

BEST AVAILABLE COPY

Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 An Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg./Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher D. Schreibe
 Searcher Phone # 308-4292
 Searcher Location CMC 6403
 Date Searcher Picked Up _____
 Date Completed 10/23
 Searcher Prep & Review Time 14
 Client Prep Time _____
 Online Time 8

Type of Search

NA Sequence (#) 2
 AA Sequence (#) _____
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel Orbit _____
 Dr. Link _____
 Lexis Nexis _____
 Sequence Systems Compuserk
 WWW Internet _____
 Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 05:44:15 ; Search time 3600 Seconds
(without alignments)
12584.318 Million cell updates/sec

Title: US-09-868-987-1
Perfect score: 1864
Sequence: 1 atggacttcgcattatgtc.....atgacaaattcagataatgc 1864

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 452938

Minimum DB seq length: 5
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_pbg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	29.6	1.6	95	9 AI349643	AI349643 ta75h07.x
2	29.6	1.6	97	29 BZ378901	BZ378901 SALK_1121
3	28.2	1.5	79	28 BH97060	BH97060 3526_1_6
C 4	28	1.5	93	12 BJ055533	BJ055533 BJ055533

C	5	27.2	1.5	90	9	AI271303	AI271303 qw69602.x
	6	27.2	1.5	99	9	AL667436	AL667436
	7	26.6	1.4	91	9	AW733398	AW733398 sk73a03.y
C	8	26.6	1.4	97	9	AU061738	AU061738
	9	26.6	1.4	98	9	AV675896	AV675896
	10	26.6	1.4	100	13	BQ795262	BQ795262 EST_4200
	11	26.4	1.4	76	28	AZ781137	AZ781137 2M0019C09
	12	26.2	1.4	99	9	AI769234	AI769234 wg35605.x
C	13	26	1.4	80	9	AA647052	AA647052 vn39a11.r
C	14	26	1.4	80	29	AL938880	AL938880 Arabidops
C	15	26	1.4	91	9	AU263883	AU263883
	16	25.8	1.4	87	28	BH861222	BH861222 SALK_0348
	17	25.8	1.4	93	29	CC326207	CC326207 RRJ093.Ba
	18	25.8	1.4	98	9	AA683513	AA683513 zF34a09.s
C	19	25.8	1.4	100	12	BI938442	BI938442 des3b09.y
	20	25.6	1.4	76	28	BH791176	BH791176 SALK_0588
	21	25.6	1.4	78	13	BU879460	BU879460 V060E04.P
	22	25.6	1.4	87	28	AZ309857	AZ309857 1M0017M12
	23	25.6	1.4	89	28	BH810836	BH810836 SALK_0512
C	24	25.6	1.4	94	12	BI692389	BI692389 603342884
	25	25.6	1.4	95	28	BH850307	BH850307 SALK_0710
C	26	25.6	1.4	100	13	BX298804	BX298804
C	27	25.4	1.4	81	9	AU013607	AU013607 AU013607
C	28	25.2	1.4	70	28	AZ789636	AZ789636 2M0037C09
C	29	25.2	1.4	79	14	CA330429	CA330429 haa98e07.
C	30	25.2	1.4	81	10	BF102955	BF102955 601646972
C	31	25.2	1.4	88	9	AU259801	AU259801 AU259801
C	32	25.2	1.4	91	29	CC459608	CC459608 SALK_1307
C	33	25.2	1.4	99	9	AA724210	AA724210 a109f07.8
C	34	25	1.3	86	28	AZ858741	AZ858741 2M0164K03
C	35	25	1.3	88	9	AA807361	AA807361 OF50d02.8
C	36	25	1.3	90	9	AI348898	AI348898 ta98603.x
C	37	25	1.3	95	28	AZ801340	AZ801340 2M0059G14
C	38	25	1.3	96	9	AV672532	AV672532 AV672532
C	39	25	1.3	96	29	AG024431	AG024431 Oryza sat
C	40	25	1.3	98	14	CA995243	CA995243 rg56b12.y
C	41	24.8	1.3	79	9	AU007905	AU007905 AU007905
C	42	24.8	1.3	88	9	AI344952	AI344952 tb01c10.x
C	43	24.8	1.3	90	9	AU268769	AU268769 AU268769
C	44	24.8	1.3	91	9	AI251961	AI251961 qv57b12.x
C	45	24.8	1.3	94	9	AI270794	AI270794 qw51a11.x

ALIGNMENTS

RESULT 1
AI349643/c
LOCUS
ta75h07.x2 NCI_CGAP_HSC2 Homo sapiens cDNA clone IMAGE:2049949 3',
DEFINITION
mRNA sequence.
ACCESSION
AI349643
VERSION
AI349643.1 GI:4086849
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 95)
AUTHORS
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 131 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source
Location/Qualifiers
1..95
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2049949"
/tissue_type="stem cell 34+/38+"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP HSC2"
/notes="Organ: bone marrow; Vector: pAMP1; mRNA made from bone marrow, stem cells 34+/38+, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 400 bp. Primary library, non-amplified."

BASE COUNT 73 a 4 c 14 g 4 t
ORIGIN

Query Match 1.6%; Score 29.6; DB 9; Length 95;
Best Local Similarity 57.8%; Pred. No. 5.9e+04;
Matches 53; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1592 GCTACAACTCTATCAGTTTGTGTAATGCTTTTGTATAGCGGCTCTCTGCTTTAAT 1651
|||||
Db 93 GATACGACITTTTTTTTTTTTTTTTTTTTTTTATCCITTTTTTTTTTTTTTTT 34
|||||

QY 1652 TTTCGATTTATTATGACCATAGGATTCCTCT 1683
|||||
Db 33 TTTTCTTTTTTTTCCCTCTCTGCTCTCT 2
|||||

RESULT 2
BZ378901
LOCUS
DEFINITION
SALK 112113.42.20.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_112113.42.20.x, genomic survey sequence.

ACCESSION BZ378901
VERSION BZ378901.1 GI:25470183
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 97)

REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab , C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmermann,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.
Class: TDNA tagged.

FEATURES
source
Location/Qualifiers
1..97
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_112113.42.20.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines

each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html

BASE COUNT 24 a 22 c 17 g 34 t
ORIGIN

Query Match 1.6%; Score 29.6; DB 29; Length 97;
Best Local Similarity 64.7%; Pred. No. 6e+04; Mismatches 0; Gaps 0;
Matches 44; Conservative 0; Indels 24; Indels 0; Gaps 0;

QY 790 AAAGCATGATTTCTTGACAGGATGCAAAAACCTTTGGCTGTTCTCGAAGTCTTTCT 849
|||||
Db 14 AACAAAGATTTCTTGAGCGAAGCAAAATGTTGGGCTCTCTTTCCGTTGATCCTCT 73
|||||

QY 850 TTTTAGGTT 857
|||||
Db 74 TTCAGGTT 81
|||||

RESULT 3
BH897060
LOCUS
DEFINITION
3526.1.6.1.F02.1EL.y.1.3526 - RescueMu Grid K Zea mays genomic, genomic survey sequence.

ACCESSION BH897060
VERSION BH897060.1 GI:22232464
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 79)

REFERENCE Walbot,V.
A Maize genomic sequences found using engineered RescueMu transposon survey sequence.
Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 3526.1.6.1 row: 6
Class: transposon-tagged.

FEATURES
source
Location/Qualifiers
1..79
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="3526 - RescueMu Grid K"
/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid K was grown at Molokai, Hawaii in Winter 2000-2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 15 a 21 c 16 g 27 t
ORIGIN


```

Query Match      1.5%; Score 28.2; DB 28; Length 79;
Best Local Similarity 64.6%; Pred. No. 1.2e+05;
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 608 TTGGCTCAGCACTCTCTTCTCTCTAGATACAGGCTATTAAAGGGTTTCCTTTGACA 667
DB 14 TTTTCTCTTCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 73

QY 668 TTGAT 672
DB 74 GGGAT 78

RESULT 4
BJ055533/c
LOCUS
DEFINITION
BJ055533
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 93)
AUTHORS
Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara,
Y.

TITLE
Expressed genes in X. laevis embryo
JOURNAL
Unpublished
CONTACT: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6856
Email: tshini@genes.nig.ac.jp.

FEATURES
Location/Qualifiers
1..93
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL009020"
/tissue_type="whole embryo"
/dev_stage="stage 15"
/clone_lib="NIBB Mochii normalized Xenopus neurula
library"

BASE COUNT      22 a 18 c 15 g 32 t 6 others
ORIGIN

Query Match      1.5%; Score 28; DB 12; Length 93;
Best Local Similarity 60.6%; Pred. No. 1.3e+05;
Matches 43; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 448 TGGGATGTTCTTGCTATGGGATGCGGTAGATGCAATGTTCTTGTTTCGAAAGAAAT 507
DB 71 TAGGCTCCTTNCCTTTATGGGATGCGGTAGATGCAATGTTCTTGTTTCGAAAGAAAT 507

QY 508 CCGAGAGGAAT 518
DB 11 CCAGCAGGAAT 1

RESULT 5
A1271303/c
LOCUS
DEFINITION
Qw68g02.x1 NCI_CGAP_Ov33 Homo sapiens CDNA clone IMAGE:1996370 3',
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
EST.

A1271303
A1271303.1 GI:3890470
EST.

```

```

SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 90)
REFERENCE
AUTHORS
TITLE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: I.M.A.G.E. Consortium, LLNL
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES
Location/Qualifiers
1..90
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1996370"
/sex="female"
/tissue_type="borderline ovarian carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Ov33"
/notes="Organ: ovary; Vector: pAMP1; mRNA made from
borderline ovarian carcinoma, CDNA made by oligo-dT
priming. Directionally cloned. Size-selected on agarose
gel, average insert size 500 bp. Primary library,
non-amplified."
57 a 9 c 13 g 3 t 8 others
BASE COUNT
ORIGIN

Query Match      1.5%; Score 27.2; DB 9; Length 90;
Best Local Similarity 53.4%; Pred. No. 1.9e+05;
Matches 47; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1596 CAACTCTATCAGTTTGTATGCTTTTATAGCGGGCTCTCTGCTTAATTTTG 1655
DB 89 CGATACAGACCTTTTGTATGCTTTTATAGCGGGCTCTCTGCTTAATTTTG 1655

QY 1656 CATTATTATGACCATAGGATTTCTCT 1683
DB 29 TTTTTCCTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2

RESULT 6
AL667436
LOCUS
DEFINITION
AL667436 directional larval CDNA library Ciona intestinalis CDNA
clone 0192E09 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
1 (bases 1 to 99)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
IMPORTANT: this sequence may contain errors. The Ciona intestinalis
library from which the clone was isolated may be contaminated with
cDNAs from bacteria or other Eukarya.

```

Directional larval cDNA library originate from Dr.M.Branno, Stazione A.Dohn, Naples, Italy, and was prepared in pBluescript2SK+.

FEATURES
source 1. .99 Location/Qualifiers

/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon.7719"
/clone="019ZE09"
/clone_lib="directional larval cDNA library"
/notes="Vector: pBluescript2SK+"
44 a 5 c 10 g 39 t 1 others

BASE COUNT
ORIGIN

Query Match 1.5%; Score 27.2; DB 9; Length 99;
Best Local Similarity 71.4%; Pred. No. 2e+05;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1730 ATGGTCGTAAGCAAAATCGCTCAAAATAGTACCGTTAAACTTAATCT 1778
Db 45 ATTTCCTGTAAGAAATTAGATAAAATTAGTAAAGTTAAGTTAAATNT 93

RESULT 7

AW733398

LOCUS AW733398 91 bp mRNA linear EST 03-DEC-2001
DEFINITION SK73a03.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl016-9701 5', mRNA sequence.

ACCESSION AW733398

VERSION AW733398.1

KEYWORDS GI:7639068

SOURCE EST.

ORGANISM Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 91)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,Y., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project

JOURNAL Unpublished

COMMENT Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63109, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert Length: 747 Std Error: 0.00

Seq primer: -40RP from Gibco

High quality sequence stop: 85.

Location/Qualifiers

1. .91

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-9701"

/tissue type="Immature flowers of field grown plants"

/lab host="Xl10-Gold"

/clone_lib="Gm-cl016"

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments directionally XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."

BASE COUNT 16 a 15 c 24 g 36 t
ORIGIN

Query Match 1.4%; Score 26.6; DB 9; Length 91;

Best Local Similarity 60.3%; Pred. No. 2.6e+05;

Matches 44; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 617 GCACCTCTTTTCTCTAGATACAGGCTATTAAAGGTTTGCCTTGACATGATTTTA 676

Db 18 GGGTGTCTTGTCTCTACCGCTGATTTATTGACGGATCGCTTTTCAACGATTTAA 77

QY 677 GGAATTTTCTCTT 689

Db 78 GGTCTTTTCTTTT 90

RESULT 8

AU061738/c

LOCUS AU061738

DEFINITION AU061738 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium

discoideum cDNA clone SLF614, mRNA sequence.

ACCESSION AU061738

VERSION AU061738.1

KEYWORDS GI:4882842

SOURCE EST.

ORGANISM Dictyostelium discoideum

Dictyostelium discoideum

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE 1 (bases 1 to 97)

AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,

Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,

Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.

Developmental cDNA in Dictyostelium discoideum

Unpublished

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp

PROJECT = Dictyostelium discoideum cDNA project in Japan.

Location/Qualifiers

1. .97

/organism="Dictyostelium discoideum"

/mol_type="mRNA"

/strains="AX4"

/db_xref="taxon:44689"

/clone="SLF614"

/dev stage="slug"

/clone_lib="Dictyostelium discoideum SL (H.Urushihara)"

BASE COUNT 48 a 7 c 6 g 35 t

ORIGIN

Query Match 1.4%; Score 26.6; DB 9; Length 97;

Best Local Similarity 59.5%; Pred. No. 2.7e+05;

Matches 44; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 647 ATTAAGGGTTTGCCTTGACATGATTTTAGGAATTTTCTTCAATGTTTACGGCTCTT 706

Db 91 ATTACTAGTTTGTGTACAAATGATTTTACAAANTTAATTTTAACTTTTAAATGTA 32

QY 707 TTCATGACTAAAT 720

Db 31 TTGATATTAAATTT 18

```

Seq primer: T7.
Location/Qualifiers
1..100
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Shiraz"
/db_xref="taxon:29760"
/db_clone="RB004A06"
/dev_stage="ripening stage"
/clone_lib="Ripening Grape berries Lambda Zap II Library"
/note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco RI
; Site 2: XhoI; Oriented library, construction described
in Generation of ESTs from Grape Berry (skin, pulp or
seeds) at various developmental stages by Ferrier,N.,
Ageorges,A., Abbal,P., Romieu,C. in J. Plant Physiol. 158
(12): 1575-83 2001"
20 a 13 c 8 g 59 t
BASE COUNT
ORIGIN

Query Match 1.4%; Score 26.6; DB 13; Length 100;
Best Local Similarity 60.3%; Pred. No. 2.7e+05;
Matches 44; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 495 TATTCGAAGAATCCGAGAGGATTTTTTATGCTCTCAAGTCTTAAAAATCTGTAGAAA 554
|||||
DB 80 TATTTTAATCAATGAAGTTGCAATTTATATGCTCTCAAAAAA
|||||
QY 555 AAGGATATACCAA 567
|||||
DB 20 AAAAAAAAAAAAA 8

RESULT 11
AZ781137
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ781137 76 bp DNA linear GSS 16-FEB-2001
2M0019C09F Mouse 10kb plasmid UUGCWM library Mus musculus genomic
clone UUGC2M0019C09 F, genomic survey sequence.
AZ781137
AZ781137.1 GI:12913527
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 76)
Dunn,D., Aiyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT,
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: rdunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0019 row: C column: 09
Seq primer: CGTTGTAAACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 76.
Location/Qualifiers
1..76
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_clone="UUGC2M0019C09"
/sex="Male"

```

/lab host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone lib="Mouse 10kb plasmid UUGCM library"
/notes="Vector: pWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (GII4732114[gb|AF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 22 a 10 c 14 g 30 t
ORIGIN
Query Match 1.4%; Score 26.4; DB 28; Length 76;
Best Local Similarity 59.2%; Pred. No. 2.8e+05;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 702 CTCCTTTCATGACTAAATTTTCTCATGCTGGAGTAAGACCCACATACAGT 761
|||||
Db 1 CTCCTTATTTTATTTTCTCTAGCTGGAGATGATGAACAGAAT 60
QY 762 TGCATATGATGAATAA 777
Db 61 CCAAGTAACCTGATTA 76

RESULT 12
AI769234
LOCUS
DEFINITION
AI769234 99 bp mRNA linear EST 19-DEC-1999
IMAGE:2367104 3', mRNA sequence.
AI769234
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 99)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1723 Std Error: 0.00
Seq primer: -400P from Gbco
High quality sequence stop: 85.
Location/Qualifiers
1. .99
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2367104"
/lab host="DH10B"
/clone lib="Soares NSF F8 9W OT PA P S1"
/notes="Organ: pooled; Vector: pWT3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.

FEATURES
source
1. .80
/organism="Mus musculus"
/mol_type="mRNA"
/strains="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:1023548"
/sex="females"
/tissue_type="whole skin"

Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NbHSP pool 1:
309384-310919, 323208-325895 Soares Nb2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 27 a 10 c 9 g 53 t
ORIGIN
Query Match 1.4%; Score 26.2; DB 9; Length 99;
Best Local Similarity 58.2%; Pred. No. 3.2e+05;
Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 621 TTCTTTTCTCTAGATACAGGCGCTATTAAAGGTTTGCTTTGACATTGATTTAGGAA 680
|||||
Db 7 TTTTCTTTTACTTTGTAAGGGGAATTTTATCAATATATCTTTTACGTTATCATTAATCT 66
QY 681 TTTTCTCTCAATGTTTAC 699
|||||
Db 67 TTTACAATTTCTAAGTTTTC 85

RESULT 13
AA647052/c
LOCUS
DEFINITION
AA647052 80 bp mRNA linear EST 28-OCT-1997
IMAGE:1023548 5', similar to gb:K02109 Mouse 3T3-L1 lipid binding
protein mRNA, complete cds (MOUSE);, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
AA647052
AA647052.1 GI:2573481
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 80)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:574324
Trace considered overall poor quality
Putative full length read
vector to vector length is 199
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .80
/organism="Mus musculus"
/mol_type="mRNA"
/strains="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:1023548"
/sex="females"
/tissue_type="whole skin"


```

BH861222      BH861222      87 bp      DNA      linear      GSS 05-AUG-2002
LOCUS          SALK_034832 Arabidopsis thaliana TDNA insertion lines Arabidopsis
DEFINITION     thaliana genomic clone SALK_034832, genomic survey sequence.
ACCESSION      BH861222
VERSION        BH861222.1 GI:22096548
KEYWORDS       GSS.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
               ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE      1 (bases 1 to 87)
AUTHORS        Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
               , C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
               , Zimmerman,J. and Ecker,J.R.
TITLE          A Sequence-Indexed Library of Insertion Mutations in the
JOURNAL        Arabidopsis Genome
COMMENT        Unpublished
               Contact: Joseph R. Ecker
               Salk Institute Genomic Analysis Laboratory (SIGNAL)
               The Salk Institute for Biological Studies
               10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
               Tel: 858 453 4100 x1752
               Fax: 858 558 6379
               Email: ecker@salk.edu
               This is single pass sequence recovered from the left border of
               TDNA.
Classes: TDNA tagged.
FEATURES             Location/Qualifiers
                     1..87
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_034832"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains one or more TDNA insertion
                     elements. The resultant fragment for each line was
                     directly sequenced to determine the genomic sequence at
                     the site of insertion. Details of the protocols used can
                     be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT        26 a 10 c 12 g 39 t
ORIGIN
Query Match      1.4%; Score 25.8; DB 28; Length 87;
Best Local Similarity 60.9%; Pred. No. 3.9e+05;
Matches 42; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY      667 ATGATTAGGAATTTCTCTCAATGTTTACGGCTCTTTTCATGACTAAATTTTCTT 726
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
        6 AATGATTGGTAACCTTAATGACGTCAATTAAGAGATTTTTCCTGATTAAATTTCTTA 65

QY      727 CATCGGTG 735
Db      |||||

QY      66 TTGGCATTG 74

RESULT 17
CC326207      CC326207      93 bp      mRNA      linear      GSS 16-MAY-2003
LOCUS          RRJ093 BayGenomics Gene Trap Library pGT2Lxf Mus musculus cDNA,
DEFINITION     mRNA sequence.
ACCESSION      CC326207
VERSION        CC326207.1 GI:30795378
KEYWORDS       GSS.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 93)
AUTHORS        BayGenomics.

```

```

http://baygenomics.ucsf.edu/
TITLE          Unpublished
JOURNAL        JOURNAL
COMMENT        Contact: BayGenomics
               Bay Area Functional Genomics Consortium (BayGenomics)
               Email: info@baygenomics.ucsf.edu
               Sequence tag generated by 5' RACE of total RNA from gene trap ES
               cell line. ES cell lines harboring insertion mutation of target
               gene are available upon request from BayGenomics. Annotation
               information available from
               http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=
               CELL_LINE&KEY=RRJ093
               Class: Gene Trap.
FEATURES             Location/Qualifiers
                     1..93
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129 Ola"
                     /db_xref="taxon:10090"
                     /sex="Male"
                     /cell_type="Embryonic stem cell"
                     /clone_lib="BayGenomics Gene Trap Library pGT2Lxf"
                     /note="Vector: pGT2Lxf"
BASE COUNT        8 a 15 c 7 g 63 t
ORIGIN
Query Match      1.4%; Score 25.8; DB 29; Length 93;
Best Local Similarity 63.9%; Pred. No. 3.9e+05;
Matches 39; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY      669 TGATTATTAGGAATTTCTCTCAATGTTTACGGCTCTTTTCATGACTAAATTTTCTTCA 728
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
        3 TGTGTTTTTTATTCGATTGTTTATTTTACGTTTTTTTCTGTCATCTTCTTCTTCT 62

QY      729 T 729
Db      63 T 63

RESULT 18
AA683513      AA683513      98 bp      mRNA      linear      EST 19-DEC-1997
LOCUS          zf34a09.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
DEFINITION     IMAGE:378808 3', mRNA sequence.
ACCESSION      AA683513
VERSION        AA683513.1 GI:2670111
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 98)
AUTHORS        Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
               Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
               ,J., Moore,B., Schellenberg,K., Steptoe,N., Tan,F., Theising,B.,
               White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE          WashU-NCI human EST Project
JOURNAL        Unpublished
COMMENT        Contact: Wilson RK
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: est@watson.wustl.edu
               This clone is available royalty-free through LLNL; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               Seq primer: -40ml3 fwd. ET from Anersham
               High quality sequence stop: 85.
FEATURES             Location/Qualifiers
                     1..98
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:378808"

```

```

/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NBH19W"
/notes="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCGAATGGAGCGCGCATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBH19W."
BASE COUNT      34 a   20 c   10 g   34 t
ORIGIN

Query Match      1.4%; Score 25.8; DB 9; Length 98;
Best Local Similarity 56.5%; Pred. No. 4e+05;
Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1383 TGTATTACGACATTTCTTTTGAAGAAAATCAATAGATTGCAAGCCATTGGTGCTT 1442
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 89 TCTTGATCCCAAGTTATTTTGAAGAAAACCTAATTGAACAGGTATGGTTGGGAGCA 30

QY 1443 TAATGACTGTTATGGGTATTCATT 1467
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 TAATAAATGTGTTTGAAGATGTT 5

RESULT 19
BI938442/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI938442
des3b09.y1 Wellcome CRC pRN3 St13 17 egg animal cap Xenopus laevis
cDNA clone IMAGE:3475528 5', mRNA sequence.
BI938442
BI938442.1 GI:16252914
EST.
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 100)
Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
Waterston,R. and Wilson,R.
Washu Xenopus EST project, 1999
Unpublished
Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Library constructed by N. Garrett, E. ellefroid, and A.M. Zorn
(Wellcome/CRC Institute). DNA Sequencing by: Washington University
Genome Sequencing Center
Clone Distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Putative full length read
vector to vector length is
Seq primer: -4ORP from Gibco.
Location/Qualifiers
1..100
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:3475528"

FEATURES
source

```

```

/tissue_type="egg, subtracted by stage 13-17 animal cap"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Wellcome CRC pRN3 St13 17 egg animal cap"
/notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieukoop and Faber. Library is subtracted
and was constructed by N. Garrett, E. Ellefroid, and A.M.
Zorn, (Wellcome/CRC Institute)."
BASE COUNT      70 a   11 c    5 g   14 t
ORIGIN

Query Match      1.4%; Score 25.8; DB 12; Length 100;
Best Local Similarity 60.9%; Pred. No. 4e+05;
Matches 42; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 656 TTTCGTTTGACATTCATTTTAGGAATTTCTTCAATGTTTACCGCTCTTTTCATGACT 715
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 TTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTCTATTTTATGGCC 31

QY 716 AAATTTTTC 724
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30 AAATGTTTC 22

RESULT 20
BH791176
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BH791176
SALK 058865.42.45.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_058865.42.45.x, genomic
survey sequence.
BH791176
BH791176.1 GI:19884747
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 76)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadriab
,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 3' end of
At5g29303.
Glass: TDNA tagged.
Location/Qualifiers
1..76
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK 058865.42.45.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"
26 a   12 c    5 g   33 t
BASE COUNT
ORIGIN

```

```

Query Match      1.4%; Score 25.6; DB 28; Length 76;
Best Local Similarity 70.8%; Pred. No. 4.2e+05;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1625 TTTATAGCGGCTCTCTGCTTTAAATTTTGCATTTATTATGACCATA 1672
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 TTTATAGCAATCTACTACTGATGATGATTTATTATTATTGACCAA 68

RESULT 21
BU879460
LOCUS      BU879460
DEFINITION BU879460 Populus flower cDNA library Populus balsamifera subsp.
trichocarpa cDNA 5 prime, mRNA sequence.
ACCESSION  BU879460
VERSION     BU879460.1 GI:24070984
KEYWORDS   EST.
SOURCE     Populus balsamifera subsp. trichocarpa
ORGANISM   Populus balsamifera subsp. trichocarpa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE  1 (bases 1 to 78)
AUTHORS   Unneberg,P., Bhalarao,R.R., Jansson,S. and Sterky,F.
TITLE     The poplar tree transcriptome: Analysis of expressed sequence tags
          from multiple libraries
JOURNAL   Unpublished
COMMENT   Contact: BHALERAO RUPALI R.
          Umea Plant Science Center
          Department of Plant Physiology
          University of Umea, 901 87 Umea, Sweden
          Tel: +46 90 786 5279
          Fax: +46 90 786 6676
          Email: rupali.bhalarao@plantphys.umu.se.

FEATURES
    source
        1..78
            /organism="Populus balsamifera subsp. trichocarpa"
            /mol_type="mRNA"
            /sub_species="trichocarpa"
            /db_xref="taxon:3694"
            /clone_lib="Populus flower cDNA library"
            /notes="Organ: flower"
BASE COUNT  30 a 9 c 14 g 25 t
ORIGIN
Query Match      1.4%; Score 25.6; DB 13; Length 78;
Best Local Similarity 59.7%; Pred. No. 4.2e+05;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 762 TGCATATGATGAATAAGTTCGTGGGATAAAGCATGATTTCTTGAGAGGATGCMAAAAC 821
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TGTCTTTAGTACTATAACGAACACTAATAAGTGAATTTAGAGAGAAAGTAACAAA 61

QY 822 TTTGGGCTGTTT 833
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 TTCCAGCTCTT 73

RESULT 22
AZ309857
LOCUS      AZ309857
DEFINITION 1M0017M12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0017M12 F, genomic survey sequence.
ACCESSION  AZ309857
VERSION     AZ309857.1 GI:10351268
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 87)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
          Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dduwn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0017 row: M column: 12
Seq primer: CGTTGTAACACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 87.
Location/Qualifiers
    1..87
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UUGC1M0017M12"
        /sex="Male"
        /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
        /clone_lib="Mouse 10kb plasmid UUGC1M library"
        /note="Vector: PWD42nv; Purified genomic DNA from M.
        musculus C57BL/6J (male) was obtained from the Jackson
        Laboratory Mouse DNA Resource
        (http://www.jax.org/resources/documents/dnares/). The DNA
        was hydrodynamically sheared by repeated passage through a
        0.005 inch orifice at constant velocity. The sheared DNA
        was blunt end-repaired with T4 DNA polymerase and T4
        polynucleotide kinase. Adaptor oligonucleotides were
        ligated to the blunt ends in high molar excess. The
        adaptor DNA was purified and size-selected for a 9.5 to
        10.5 kb range using preparative agarose gel
        electrophoresis. Vector DNA was prepared from a derivative
        of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
        inducible derivative of plasmid R1. The vector was ligated
        with adaptors complementary to the insert adaptors and
        purified. The sheared, adaptor mouse DNA was annealed to
        adaptor vector DNA, and transformed into
        chemically-competent E. coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."
BASE COUNT  17 a 8 c 20 g 42 t
ORIGIN
Query Match      1.4%; Score 25.6; DB 28; Length 87;
Best Local Similarity 59.7%; Pred. No. 4.3e+05;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 674 TTAGGAATTTCTCTTCAATGTTTACGGCTCTTTTCATGACTAAATTTTCTTCATGCTG 733
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTGTGTGTGTTAAATGAATTTTACTTTTCAATTTATGTTGATGGTCTCTGCTTATG 60

QY 734 TGGATGAATAAG 745
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TGTATGTATATG 72

RESULT 23
BH810836
LOCUS      BH810836
DEFINITION SALK 051270 Arabidopsis thaliana TDNA insertion lines Arabidopsis
thaliana genomic clone SALK_051270, genomic survey sequence.
ACCESSION  BH810836
VERSION     BH810836.1 GI:20388654
KEYWORDS   GSS.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana

```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids I (bases 1 to 89)

REFERENCE

AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.

TITLE

A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL

COMMENT Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA. This sequence lies within an annotated intron of At4g37610.

Class: TDNA tagged.

FEATURES

source

1..89

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK 051270"

/clone_lib="Arabidopsis thaliana TDNA insertion lines"

/note="PCR was performed on Arabidopsis thaliana lines

each of which contains one or more TDNA insertion

elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at

the site of insertion. Details of the protocols used can

be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT

29 a

4 c

16 g

33 t

7 others

Query Match

Best Local Similarity 1.4%; Score 25.6; DB 28; Length 89;

Matches 43; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY

1737 GTAAGAAATCGCTCAAAATAGTACCGTTAACTTAATCTAACGTGTAGCAATATAA 1796

Db

5 GTAAAGAAAGTGTGAAGATTGTGATATAATCTTAANTTCTGATATGACAAATTTAA 64

QY

1797 AATCTCTTTGGAGCTTTA 1815

Db

65 AAGTCTTTTGAANTTA 83

RESULT 24

BI692389/c

LOCUS

DEFINITION 603342884F1 NCI_CGAP_Mam2 94 bp mRNA linear EST 18-SEP-2001

mus musculus cDNA clone IMAGE:5370669 5',

mRNA sequence.

ACCESSION

BI692389

VERSION

BI692389.1 GI:15655018

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 94)

NIH-MGC <http://mgs.nci.nih.gov/>

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11943 row: 9 column: 22

High quality sequence start: 8

High quality sequence stop: 93.

FEATURES

source

1..94

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:5370669"

/tissue_type="tumor, biopsy sample"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI CGAP Mam2"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 56 a

3 c

7 g

28 t

Query Match

Best Local Similarity 1.4%; Score 25.6; DB 12; Length 94;

Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY

668 TTGATTTAGGAATTTCTCTTCATGTTTACGGCTCTTTTCATGACTAAATTTCTTC 727

Db

88° TTTTCTTTTGCATTAATCTCCATATTTTATTTTATTTTACACATTATTTA 29

QY

728 ATGCTGTGGATGAATAAGAC 747

Db

28 ATTTTGTATATATTATAC 9

RESULT 25

BI692389/c

LOCUS

DEFINITION

BI692389

VERSION

BI692389.1 GI:21421178

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

I (bases 1 to 95)

REFERENCE

AUTHORS

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab

,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.

Zimmerman,J. and Ecker,J.R.

TITLE

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

JOURNAL

COMMENT

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA.

Class: TDNA tagged.

FEATURES

source

1..95

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

```

/clone="SALK 071085.29.45.x"
/clone lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT      29 a   16 c   10 g   40 t
ORIGIN
Query Match      1.4%; Score 25.6; DB 28; Length 95;
Best Local Similarity 55.7%; Pred. No. 4.4e+05;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 480 ATGCAATGTTCTGATTTCGAAAGAAATCCGAGAGAAATTTTATTTGTTCTCAAGTCTTA 539
    |||||
Db 94 ATTCGAGTAGTTATGATTTTGATAAAAGAAAGGACGTTGTTCCCTCATGATCCT 35
    |||||

QY 540 AAAAATCTGAGAAAAGGATATACCAA 567
    |||||
Db 34 AAAATGTCATAAAATGAAAAGGCAA 7
    |||||

RESULT 26
BX298804
LOCUS
DEFINITION
BX298804
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 100)
Govoroun,M., Guiguen,Y. and Le Gac,F.
Construction and primary characterization of normalized cDNA
libraries in rainbow trout, Oncorhynchus mykiss
Unpublished
Contact: Guiguen Y
INRA - SCRIBE
Campus de Beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at sigenasupport@jouy.inra.fr to obtain the Chromatogram of this
sequence.
Plate: 0004 row: o column: 6
Seq primer: M13R.
FEATURES
Source
1..100
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="tcay0004b.o.06"
/tissue type="adipose tissue, blood, brain,
differentiating gonads, gills, interrenal, intestine,
kidney, liver, muscle, ovary, pituitary, testis"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone lib="AGENAE Rainbow trout multi-tissues subtracted
library (tcay)"
/notes="vector: pT73D-pac; Clone distribution : AGENAE
Resource centre: Francois PIUMI,
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (LREG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, FRANCE"
BASE COUNT      45 a   7 c   23 g   25 t

```

ORIGIN

```

Query Match      1.4%; Score 25.6; DB 13; Length 100;
Best Local Similarity 66.1%; Pred. No. 4.4e+05;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 490 TCTTTATTCGAAGAATCCGAGAGAAATTTTATTTGTTCTCAAGTCTTAAAAAAT 545
    |||||
Db 1 TCAAGTACAGGAAGAACATGACAGGTATGTCAGTTGTATAAAATGTTTAAAAAAT 56
    |||||

RESULT 27
AU013607/c
LOCUS
DEFINITION
AU013607 Schizosaccharomyces pombe late log phase cDNA
Schizosaccharomyces pombe cDNA clone spc03401, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Schizosaccharomyces pombe (fission yeast)
Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomyces.
1 (bases 1 to 81)
Morimyo,M. and Mita,K.
Identification of expressed sequence tags of Schizosaccharomyces
pombe
Unpublished
Contact: Mitsuoki Morimyo
Genome Research Group
National Institute of Radiological Sciences
9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
Email: morimyo@nirs.go.jp.
Location/Qualifiers
1..81
/organism="Schizosaccharomyces pombe"
/mol_type="mRNA"
/db_xref="taxon:4896"
/strain="972"
/clone="spc08401"
/sex="h minus"
/notes="Vector: M13mp19; The cDNA library of
Schizosaccharomyces pombe was prepared by cloning cDNA
into the SmaI site of M13mp19 DNA and the direction of DNA
sequences was not always from 5' to 3'. The cDNA data of
Schizosaccharomyces pombe are available for searching on
the World Wide Web. (URL, http://www.nirs.go.jp)"
BASE COUNT      30 a   11 c   9 g   31 t
ORIGIN
Query Match      1.4%; Score 25.4; DB 9; Length 81;
Best Local Similarity 58.7%; Pred. No. 4.7e+05;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 703 TCTTTTCATGACTAAATTTTCTTCATGCTGGGATGATAGACCCACATACAGTT 762
    |||||
Db 80 TTTTATCATGTTAAATTTTATTGATAAAATCGTTATGACCAAGACGTAAAAAATCCGTT 21
    |||||

QY 763 GCATATGATGATATAA 777
    |||||
Db 20 TCATATGATGATGAA 6
    |||||

RESULT 28
AZ789636/c
LOCUS
DEFINITION
AZ789636
ACCESSION
VERSION
KEYWORDS
GSS.
70 bp DNA linear GSS 16-FEB-2001
2M0037C09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
Clone UUGC2M0037C09 R, genomic survey sequence.
AZ789636
AZ789636.1 GI:12930665
GSS.

```

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 70)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 308 Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0037 Row: C Column: 09
 Seq primer: CACACAGGAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 70.
 Location/Qualifiers
 1..70
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0037C09"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UGCLM library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (G14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 BASE COUNT 20 a 25 c 15 g 10 t
 ORIGIN
 Query Match 1.4%; Score 25.2; DB 28; Length 70;
 Best Local Similarity 62.9%; Pred. No. 5.1e+05;
 Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 625 TTCTCTCTAGATACAGGCGCTTAAAGGGTTTGCTTGACATTCATTTAGGAATTTT 684
 Db 64 TTGCCCCCTGACTGCAGTGCCTAGCGGGCTTCTGATGGTATTTTGATAGTGG 5
 QY 685 CT 686
 Db 4 CT 3
 RESULT 29
 CA330429 79 bp mRNA linear EST 04-NOV-2002
 LOCUS

DEFINITION haa98e07.y1 Fugu hgmpJ adult kidney Takifugu rubripes cDNA clone
 6362197.5', mRNA sequence.
 ACCESSION CA330429
 VERSION CA330429.1 GI:24548527
 EST:
 SOURCE Takifugu rubripes (Fugu rubripes)
 ORGANISM Takifugu rubripes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Takifugu.
 REFERENCE 1 (bases 1 to 79)
 AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy,
 S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood,
 K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
 Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,
 Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
 and Wilson, R.
 TITLE WashU Zebrafish EST Project 1998
 JOURNAL Unpublished
 COMMENT Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@watson.wustl.edu
 Library materials and construction by Greg Elgar (UK MRC HGMP-RC).
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Fugu clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Putative full length read
 The vector to vector length is 80
 Seq primer: T3 ET from Amersham.
 Location/Qualifiers
 1..79
 /organism="Takifugu rubripes"
 /mol_type="mRNA"
 /db_xref="taxon:31033"
 /clone="6362197"
 /dev_stage="adult"
 /lab_host="DH10B (T1-resistant)"
 /clone_lib="Fugu hgmpJ adult kidney"
 /note="Organ: kidney; Vector: pBluescript II KS
 (Stratagene); Site 1: XhoI; Site 2: EcoRI; Library is
 oligo-dT primed and directionally cloned using cDNA
 prepared with the Stratagene cDNA synthesis kit. Tissue
 source: adult kidney. Library materials and construction
 by Greg Elgar (UK MRC HGMP-RC)."
 BASE COUNT 30 a 12 c 12 g 25 t
 ORIGIN
 Query Match 1.4%; Score 25.2; DB 14; Length 79;
 Best Local Similarity 60.0%; Pred. No. 5.2e+05;
 Matches 42; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 486 ATGTTCTTGTATTCGAAGATCCGAGAGGATTTTATTTGTCCTCAAGCTCTTAAARAT 545
 Db 10 ATCTCTCTGCTTTGTAAATAATCTGTGAATAAATATTTTGTCTCAGCGGTTTAAAAA 69
 QY 546 CTGTAGAAAA 555
 Db 70 AAAAAAAAAA 79
 RESULT 30
 BF102955/c
 LOCUS 601646972F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4073344 5',
 DEFINITION mRNA sequence.
 ACCESSION BF102955
 VERSION BF102955.1 GI:10885481
 KEYWORDS EST.

the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html

BASE COUNT
ORIGIN

Query Match 1.4%; Score 25.2; DB 29; Length 91;
Best Local Similarity 71.7%; Pred. No. 5.3e+05;
Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
40 a 8 c 8 g 35 t

QY 1056 TCAAAATCTATTTAGTATAGCTTTAAGCTTATAGCTATACAGAT 1101

Db 13 TTAATAATTATGTCGTATATAGATTAAATATCTATACAAT 58

RESULT 33
AA724210/c

LOCUS AA724210 99 bp mRNA linear EST 31-DEC-1998
DEFINITION ai09f07.s1 Soares parathyroid_tumor_NBHPA Homo sapiens cDNA clone
1342309 3', mRNA sequence.

ACCESSION AA724210

VERSION 1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 99)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

UNPUBLISHED

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.

cDNA library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1236 Std Error: 0.00

Seq primer: -40m3 fwd. ET from Amersham

High quality sequence stop: 88.

Location/Qualifiers

FEATURES

source

1. .99

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="1342309"

/tissue_type="parathyroid tumor"

/dev_stage="adult"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares parathyroid_tumor_NBHPA"

/note="Organ: parathyroid gland; Vector: pT73D (Pharmacia
RI) with a modified polylinker; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer

[5'-TGTTACCAATCTGAAGTGGAGCGCGCCACCAATTTTTTTTTTTTTTTT
TTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT73
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."

BASE COUNT 30 a 14 c 16 g 39 t

ORIGIN

Query Match

Best Local Similarity 66.7%; Score 25.2; DB 9; Length 99;

Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1372 CTGTGCAGCTCTGTTTATAGCACATTTCTTTTGAAGAAATTCAAATAGATT 1425

Db 76 CTGCATAGTCAGCCTTATACTTGATTCTTTTGAATCAATAAATAAAATT 23

RESULT 34

AZ858741

LOCUS

DEFINITION

2M0164K03F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M0164K03 F, genomic survey sequence.

ACCESSION AZ858741

VERSION AZ858741.1

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 86)

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

UNPUBLISHED

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 309, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0164 row: K column: 03

Seq primer: CTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 86.

Location/Qualifiers

1. 86

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0164K03"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, P-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gi|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 23 a 12 c 29 g 22 t

ORIGIN

Query Match

Best Local Similarity 1.3%; Score 25; DB 28; Length 86;

Matches 64.9%; Pred. No. 5.8e+05;

Plate: 0059 row: G column: 14
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 95.

FEATURES

source
 1. .95
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="JUGC2M0059C14"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI:4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 24 a 11 c 32 g 28 t
 ORIGIN
 Query Match 1.3%; Score 25; DB 28; Length 95;
 Best Local Similarity 64.9%; Pred. No. 5.9e+05;
 Matches 37; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 143 GCAGTGTCTCAGGAAATTTACCCACCGTAGTGGAGCAAACTCCCTCAGATTTA 199
 DB 58 GCAAGCCTCACAGATATCTACCAACTTAGATTAGGCACCTCCCTCATATTAA 2

RESULT 38
 AV672532
 LOCUS
 DEFINITION
 AV672532 Nori Satoh unpublished cDNA library Ciona intestinalis
 cDNA clone citb512 5', mRNA sequence.
 ACCESSION
 AV672532 GI:10110531
 KEYWORDS
 EST.
 SOURCE
 Ciona intestinalis
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.
 REFERENCE
 1 (bases 1 to 96)
 Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
 Expressed genes in Ciona intestinalis
 Unpublished
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1. .96
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"

FEATURES

source
 1. .96
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"

/clone="citb512"
 /tissue type="whole animal"
 /dev_stage="tailbud"
 /clone_lib="Nori Satoh unpublished cDNA library"
 BASE COUNT 39 a 5 c 12 g 40 t
 ORIGIN

Query Match 1.3%; Score 25; DB 9; Length 96;
 Best Local Similarity 69.4%; Pred. No. 5.9e+05;
 Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 1730 ATGTCGCTTAAGAAATCGTCAAAATAGTACCGTTAAACTTAATCT 1778
 DB 13 APTTCTGTGAAGAAATAGATGAATTAAGTAAAGTTAAGTTAAATGT 61

RESULT 39

AG024431
 LOCUS
 DEFINITION
 AG024431.1 GI:7683095
 ORYZA SATIVA (japonica cultivar-group)
 3' flanking sequence of Tos17 insertion in rice strain ND4563,
 genomic survey sequence.
 ACCESSION
 AG024431
 VERSION
 GSS.
 SOURCE
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1
 Miyao,A. and Hirochika,H.
 Rice insertion mutants
 Unpublished
 2 (bases 1 to 96)
 Miyao,A., Tanaka,K. and Hirochika,H.
 Direct Submission
 Submitted (25-OCT-1999) Akio Miyao, National Institute of
 Agricultural Sciences, Molecular Genetics; 2-1-2, Kannondai,
 Tsukuba, Ibaraki 305-8602, Japan (E-mail:miyao@affrc.go.jp,
 URL:http://tos.nias.affrc.go.jp/, Tel:81-298-38-7020,
 Fax:81-298-38-7020)

FEATURES

source
 1. .96
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /strain="ND4563"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="ND4563 0 101 1A"
 /clone_lib="PCR product directly amplified from rice
 genomic DNA"
 /note="the 3' end of retrotransposon Tos17 was found
 immediately upstream of this sequence."
 BASE COUNT 26 a 17 c 14 g 37 t
 ORIGIN

Query Match 1.3%; Score 25; DB 29; Length 96;
 Best Local Similarity 58.9%; Pred. No. 5.9e+05;
 Matches 43; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
 QY 327 TGATGAGCGTATATTATAGATTGGAGGCGTCATCGCTTCGGGAGCTGTTCTTCTGAATC 386
 DB 1 TGATCATAGTATGTTACGAACCTGGATCAACCATCAGTATCAAAATTTGTCTAATTCATA 60

QY 387 TTTTGCTTAATCTG 399
 DB 61 TTTTGCTTTCTG 73

RESULT 40

CA995243
 LOCUS
 CA995243 98 bp mRNA linear EST 07-JAN-2003

```

DEFINITION   rg26b12.y1 Meloidogyne hapla J2 SL1 TOPO v1 Meloidogyne hapla cDNA
              5', mRNA sequence.
ACCESSION    CA995243
VERSION      CA995243.1 GI:27540114
KEYWORDS     EST.
SOURCE       Meloidogyne hapla
ORGANISM     Meloidogyne hapla
              Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
              Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE    1 (bases 1 to 98)
AUTHORS      McCarter,J., Clifford,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
              Dante,M., Maier,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
              Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.,
              Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,
              M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
              Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
              Wilson,R.
TITLE        The Washington Univ. Nematode EST Project, 1999
JOURNAL      Unpublished
COMMENT      Contact: McCarter JP
              The Washington Univ. Nematode EST Project, 1999
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              The library was constructed by Claire Murphy and Dr. James McCarter
              at Washington University, St. Louis. J2 were provided by Dr.
              Valerie Williamson of the University of California at Davis
              (vmwilliamson@ucdavis.edu).
              Putative full length read
              The vector to vector length is 99
              Seq primer: -40RP from Gibco.
              Location/Qualifiers
FEATURES             1..98
                     /organism="Meloidogyne hapla"
                     /mol_type="mRNA"
                     /db_xref="taxon:6305"
                     /seq_start="J2"
                     /lab_host="DH10B"
                     /clone_lib="Meloidogyne hapla J2 SL1 TOPO v1"
                     /notes="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI;
                     Site 2: EcoRI; The library was constructed by Claire
                     Murphy and Dr. James McCarter at Washington University,
                     St. Louis. Oligo(dT)-SL1 PCR based library. Meloidogyne
                     hapla J2 cDNA PCR products of size >400 nucleotides
                     containing SL1 on the 5' end and oligo(dT) on the 3' end
                     were non-directionally cloned into pCRII-TOPO(Invitrogen)
                     following the TOPO TA cloning protocol. J2 were provided
                     by Dr. Valerie Williamson of University of California at
                     Davis (vmwilliamson@ucdavis.edu)."
BASE COUNT      41 a 11 c 9 g 37 t
ORIGIN
Query Match      1.3%; Score 25; DB 14; Length 98;
Best Local Similarity 55.1%; Pred. No. 5.9e+05;
Matches 49; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1385 TTTATGACATCTTTCTTTGAAGAAAATCAATAGATTGCAAGCCATTGGTGCTTTA 1444
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 TATATAACTAATCGTTTGGTTTATTTTAAATAAATAATTTTCTTCAATGGCTATATCA 66
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1445 ATGACTGTATGGGTGATTTCATTAAACAA 1473
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ATTTGTACATTAAAGTAACTTTTAAAAA 95

RESULT 41
AU007905/c
LOCUS       AU007905 Schizosaccharomyces pombe late log phase cDNA
DEFINITION  Schizosaccharomyces pombe cDNA clone spc02639, mRNA sequence.
ACCESSION   AU007905

```

```

VERSION       AU007905.1 GI:3344363
KEYWORDS      EST.
SOURCE        Schizosaccharomyces pombe (fission yeast)
ORGANISM      Schizosaccharomyces pombe
              Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
              Schizosaccharomycetales; Schizosaccharomycetaceae;
              Schizosaccharomycetes.
REFERENCE    1 (bases 1 to 79)
AUTHORS      Morimyo,M. and Mita,K.
TITLE        Identification of expressed sequence tags of Schizosaccharomyces
              pombe
JOURNAL      Unpublished
COMMENT      Contact: Mitsuo Morimyo
              Genome Research Group
              National Institute of Radiological Sciences
              9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
              Email: morimyo@nirs.go.jp.
              Location/Qualifiers
FEATURES             1..79
                     /organism="Schizosaccharomyces pombe"
                     /mol_type="mRNA"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc02639"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web. (URL, http://www.nirs.go.jp)"
BASE COUNT      40 a 5 c 13 g 20 t
ORIGIN
Query Match      1.3%; Score 24.8; DB 9; Length 79;
Best Local Similarity 59.4%; Pred. No. 6.3e+05;
Matches 41; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1663 TATGACCATAGGATCTCTTAGGAACCTTATCGTCTCTTATATTCACACCTCTGTT 1722
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 TATGCTATGTTCTTTTAAACAGTACTTTTAACTTTGTTTATCTGTTTACITTTAT 12
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1723 GTTGTTTAT 1731
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
11 CTAATTTAT 3

RESULT 42
AU007905/c
LOCUS       AU007905 Schizosaccharomyces pombe late log phase cDNA
DEFINITION  Schizosaccharomyces pombe cDNA clone spc02639, mRNA sequence.
ACCESSION    AU007905
VERSION      AU007905.1 GI:4082158
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 88)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              CDNA Library Preparation: David B. Krizman, Ph.D.
              CDNA Library Arrayed by: I.M.A.G.E. Consortium, LNL
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              www-bio.lnl.gov/bbrp/image/image.html

```


REFERENCE 1 (bases 1 to 94)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbsr@mail.nih.gov
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: I.M.A.G.E. Consortium, LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES
Location/Qualifiers
1..94
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1994588"
/sex="female"
/tissue_type="breast carcinoma in situ"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Br13"
/notes="Organ: breast; Vector: pAMP1; mRNA made from breast
carcinoma, cDNA made by oligo-dr priming. Directionally
cloned. Size-selected on agarose gel, average insert
size 500 bp. Primary library, non-amplified."
BASE COUNT 65 a 6 c 19 g 4 t

ORIGIN

Query Match 1.3%; Score 24.8; DB 9; Length 94;
Best Local Similarity 54.3%; Pred. No. 6.5e+05;
Matches 50; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 1592 GCTACACTCTATCAGTTTGTAAATGCTTTGTTTATAGGGGCTCCTCTCTTAAAT 1651
Db 93 GATACGACTTTTTTTTTTTTTTTTTTTTTTTTTTTGTCTTTCTCTTCTTATC 34
QY 1652 TTTCATTATTATGACCATAGGGATTCTTCT 1683
Db 33 TTTCCTTTTGTCCCTCTCTCTCTCT 2

Search completed: October 23, 2003, 09:20:09
Job time : 3604 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 09:20:14 ; Search time 3602 Seconds
(without alignments)

12577.330 Million cell updates/sec

Title: US-09-868-987-1

Perfect score: 1864

Sequence: 1 atggacttcgcgattgtc.....atgacaattcagataatgc 1864

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 71794

Minimum DB seq length: 10

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gssi:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.4	1.2	36	28	BH907371
2	19.4	1.0	34	9	AI041455
3	19.4	1.0	39	28	AZ779322
4	19.4	1.0	40	9	AI318679

C	5	19.2	1.0	40	9	AA068300
C	6	18.8	1.0	35	28	AZ785768
C	7	18.8	1.0	40	9	AU264483
C	8	18.6	1.0	38	29	BZ761950
C	9	18.4	1.0	39	12	BZ761950
C	10	18.4	1.0	35	29	BZ358798
C	11	18.2	1.0	39	14	CO1981
C	12	18.2	1.0	40	9	AI022013
C	13	18	1.0	37	9	AL849593
C	14	18	1.0	39	9	AU266702
C	15	18	1.0	39	9	AV853005
C	16	18	1.0	39	29	BZ761743
C	17	17.8	1.0	36	28	AZ647867
C	18	17.8	1.0	37	28	AZ448035
C	19	17.8	1.0	38	14	H57900
C	20	17.8	1.0	38	28	BH864598
C	21	17.8	1.0	39	28	AG227568
C	22	17.8	1.0	40	28	AZ767521
C	23	17.6	0.9	34	9	AU259110
C	24	17.6	0.9	34	29	AG218655
C	25	17.6	0.9	35	29	BZ763041
C	26	17.6	0.9	37	29	BZ768062
C	27	17.6	0.9	40	28	BH789828
C	28	17.6	0.9	40	29	BZ663364
C	29	17.4	0.9	28	9	AA863634
C	30	17.4	0.9	35	28	BH866003
C	31	17.4	0.9	38	28	BH864481
C	32	17.4	0.9	39	28	AZ772005
C	33	17.4	0.9	40	28	AZ345478
C	34	17.2	0.9	40	28	BH864835
C	35	17.2	0.9	31	9	AU265359
C	36	17.2	0.9	31	9	AA989540
C	37	17.2	0.9	33	14	CO0007
C	38	17.2	0.9	36	28	AZ481068
C	39	17.2	0.9	38	28	AZ311715
C	40	17.2	0.9	38	28	AZ583903
C	41	17.2	0.9	39	28	CC057135
C	42	17.2	0.9	39	29	AL1752023
C	43	17.2	0.9	40	9	AA116378
C	44	17.2	0.9	40	9	AJ237274
C	45	17.2	0.9	40	12	BI693481

ALIGNMENTS

RESULT 1

BH907371

LOCUS

DEFINITION

Arabidopsis thaliana genomic clone SALK_042003.18.65.x, genomic

survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH907371 36 bp DNA linear GSS 04-SEP-2002

SALK_042003.18.65.x Arabidopsis thaliana TDNA insertion lines

Arabidopsis thaliana genomic clone SALK_042003.18.65.x, genomic

survey sequence.

BH907371

BH907371.1 GI:22720304

GSS.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 36)

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab

, C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.

, Zimmerman,J. and Ecker,J.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.
Location/Qualifiers
1. .36
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_042003.18.65.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
13 t

BASE COUNT 14 a 4 c 5 g
ORIGIN

Query Match 1.2%; Score 22.4; DB 28; Length 36;
Best Local Similarity 81.2%; Pred. No. 1.9e+06;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1027 TATTCACACATTGGATCTTCAGAAAGATCA 1058
|||||
Db 5 TATTCACACACTAGTATTTTCAGAGAGATTA 36
|||||

RESULT 2
AI041455
LOCUS
DEFINITION
Ow36c10.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone IMAGE:1648914 3' similar to SW:NU4M HUMAN P03905 NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4 ; mRNA sequence.

ACCESSION AI041455
VERSION AI041455.1 GI:3280649
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 34)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 532 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .34
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1648914"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_parathyroid_tumor_NbHPA"
/notes="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia

FEATURES
source

) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer.
15'-IGTTACCAATCTGAAGTGGAGCGCGCCACCAATTTTTTTTTTTTTTTT
TTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA from sporadic parathyroid adenomas was kindly provided by Dr. Stephen Marx, National Institute of Diabetes and Digestive and Kidney Diseases, NIH."

BASE COUNT 17 a 5 c 8 g 4 t
ORIGIN

Query Match 1.0%; Score 19.4; DB 9; Length 34;
Best Local Similarity 79.3%; Pred. No. 8.2e+06;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 541 AAAATCTGTAGAAAAGGATATACCAAGG 569
|||||
Db 1 AAAATATGCCCAAAAGGATATACCAAGG 29
|||||

RESULT 3
AZ779322 39 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0015A11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0015A11 R, genomic survey sequence.

ACCESSION AZ779322
VERSION AZ779322.1 GI:12909859
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 39)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: rdunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0015 row: A column: 11
Seq primer: CACACAGGAACACAGCTATCACC
Class: plasmid ends
High quality sequence stop: 39.
Location/Qualifiers
1. .39
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0015A11"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus CS7BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a

FEATURES
source

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

Query Match 1.0%; Score 19.4; DB 28; Length 39;
Best Local Similarity 70.3%; Pred. No. 8.3e+06;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 296 GCATGCTGGTGGTGGCAATGCTATTGTTTCATGA 332
Db 1 GCTTGCTTTCATGCAATGCTATTGTTTCATGA 37

RESULT 4
LOCUS AI1318679/c 40 bp mRNA linear EST 18-DEC-1998
DEFINITION ta49g08.x1 NCI CGAP Lu25 Homo sapiens cDNA clone IMAGE:2047454 3', similar to TR:Q34096 Q34096 MURF2 PROTEIN. ;, mRNA sequence.

ACCESSION AI1318679
VERSION AI1318679.1 GI:4034614
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 40)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: I.M.A.G.E. Consortium, LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -400P from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 40
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2047454"
/tissue_type="bronchioalveolar carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu25"
/notes="Organ: lung; Vector: pAMP1; mRNA made from lung carcinoma tissue, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified."

BASE COUNT
ORIGIN

6 a 3 c 6 g 25 t

Query Match 1.0%; Score 19.4; DB 9; Length 40;
Best Local Similarity 70.3%; Pred. No. 8.4e+06;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1229 AAACATATCGAGAAATCGTTATCAGCGGACCATCG 1265
Db 40 AAACAAACAAAATAATACATATCAGGTAAACATTTG 4

RESULT 5
LOCUS AA068300/c 40 bp mRNA linear EST 07-FEB-1997
DEFINITION mm61h03.r1 Stratagene mouse embryonic carcinoma (#937317) Mus musculus cDNA clone IMAGE:532949 5', similar to TR:G1050752 KYNURENINE/ALPHA-AMINOADIPATE AMINOTRANSFERASE ;, mRNA sequence.

ACCESSION AA068300
VERSION AA068300.1 GI:1566558
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 40)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:319885
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 13.
Location/Qualifiers
1. 40
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:532949"
/tissue_type="carcinoma"
/dev_stage="embryonic"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse embryonic carcinoma (#937317)"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. P19 cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."

FEATURES
source

BASE COUNT
ORIGIN

12 a 6 c 10 g 12 t

Query Match 1.0%; Score 19.2; DB 9; Length 40;
Best Local Similarity 75.0%; Pred. No. 9.2e+06;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 921 CCTTTACCTTTTAAATCCAAAGAGCATGCGATC 952
Db 39 CCTTTAACTTTTAAATCCATAGAACATGCCAGC 8

RESULT 6
AZ785768/c

AZ785768 35 bp DNA linear GSS 16-FEB-2001
 LOCUS 1M0030C04F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M0030C04 F, genomic survey sequence.
 ACCESSION AZ785768
 VERSION GSS.
 SOURCE AZ785768.1 GI:112922858
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 35)
 DUNN,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 AUTHOR ISLAM,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weis,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 Plasmid inserts
 JOURNAL Unpublished
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0030 row: C column: 04
 Seq primer: CGTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 35.
 Location/Qualifiers
 1..35
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0030C04"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 6 a 3 c 7 g 19 t

FEATURES source

BASE COUNT 6 a 3 c 7 g 19 t
 ORIGIN
 Query Match 1.0%; Score 18.8; DB 28; Length 35;
 Best Local Similarity 76.7%; Pred. No. 1.1e+07;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 539 AAAAAATCTGTAGAAAAGATATACCAAG 568
 |||||
 Db 30 AAAAAATCTATTGAGACATAAATACCAAG 1
 |||||
 RESULT 7

AU264483/c
 LOCUS AU264483 40 bp mRNA linear EST 10-MAY-2002
 DEFINITION VS Dictyostelium discoideum CDNA clone VSD733 5', mRNA
 sequence.
 ACCESSION AU264483
 VERSION AU264483.1 GI:20523281
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 1 (bases 1 to 40)
 URUSHIHARA,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
 AUTHOR TAKEUCHI,I., Kohara,Y. and Tanaka,Y.
 Population analysis of cDNAs from unicellular and multicellular
 stages of Dictyostelium discoideum
 Unpublished
 JOURNAL Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
 Tel: 81-298-53-4664
 Fax: 81-298-53-6614
 Email: hideko@biol.tsukuba.ac.jp.
 Location/Qualifiers
 1..40
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="VSD733"
 /sex="mat A"
 /dev_stage="vegetative"
 /clone_lib="VS"
 17 a 4 c 0 g 19 t
 BASE COUNT
 ORIGIN
 Query Match 1.0%; Score 18.8; DB 9; Length 40;
 Best Local Similarity 68.4%; Pred. No. 1.1e+07;
 Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 520 TTATTGCTCAAGCTTTAAAAATCTGTAGAAAAG 557
 |||||
 Db 40 TTTTITTTTAAAAATGTTTAAAAAAGATTAAAAAG 3
 |||||
 RESULT 8
 BZ761950 38 bp DNA linear GSS 13-MAR-2003
 LOCUS SALK 083640.22.00.n Arabidopsis thaliana TDNA insertion lines
 DEFINITION Arabidopsis thaliana genomic clone SALK_083640.22.00.n, genomic
 survey sequence.
 ACCESSION BZ761950
 VERSION BZ761950.1 GI:28934503
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 38)
 ALONSO,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
 AUTHOR C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
 Zimmerman,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished
 JOURNAL Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu

This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.

FEATURES

Location/Qualifiers

1..38
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_083640.22.00.n"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 14 a 5 c 3 g 16 t

ORIGIN

Query Match 1.0%; Score 18.6; DB 29; Length 38;
Best Local Similarity 72.7%; Pred. No. 1.2e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1381 CTTGTTTATAGCACATTCTTTTGAAGAAAT 1413

|||||
Db 1 CTTTATTACACATATTTCATTGAAGATCAT 33

RESULT 9

BJ057535

LOCUS

DEFINITION BJ057535 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL104fil 5', mRNA sequence.

ACCESSION BJ057535

VERSION BJ057535.1

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 39)

AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara Y.

TITLE Expressed genes in X. laevis embryo

JOURNAL Unpublished

COMMENT

Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES

Location/Qualifiers

1..39
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL104fil"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"

BASE COUNT 11 a 2 c 4 g 21 t 1 others

ORIGIN

Query Match 1.0%; Score 18.4; DB 12; Length 39;
Best Local Similarity 67.6%; Pred. No. 1.4e+07;
Matches 25; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 509 CGAGAGGAATTTTATGTCCTCAAGTCTTAAAAAT 545

|||||
|||||

Db

RESULT 10

BJ058798/c

LOCUS

DEFINITION

Arabidopsis thaliana genomic clone SALK_133223.42.50.x, genomic survey sequence.

ACCESSION BJ058798

VERSION BJ058798.1

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (chale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

euroids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 35)

AUTHORS

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab

,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.

, Zimmerman,J. and Ecker,J.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished

COMMENT

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 X1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA.

Class: TDNA tagged.

Location/Qualifiers

1..35

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_133223.42.50.x"

/clone_lib="Arabidopsis thaliana TDNA insertion lines"

/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 15 a 5 c 6 g 9 t

ORIGIN

Query Match 1.0%; Score 18.2; DB 29; Length 35;

Best Local Similarity 74.2%; Pred. No. 1.5e+07;

Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1466 TTAAACAATACCTTCATCATTTTCATCGTA 1496

|||||

Db 32 TTATCCATATTTTCACCATGTTTGTAGTA 2

|||||

RESULT 11

CO1981/c

LOCUS

DEFINITION

HUMS0003999 Human adult (K.Okubo) Homo sapiens cDNA, mRNA

sequence.

ACCESSION CO1981

VERSION CO1981.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 39)

AUTHORS

TITLE

JOURNAL

COMMENT

Okubo,K.
BodyMap: human gene expression database
Unpublished
Contact: Okubo,K.
Institute for Molecular and Cellular Biol
Osaka University
1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan
Tel: 06-877-5111(ex.3315)
Email: kouseku@imcb.osaka-u.ac.jp
We are not submitting the same cDNA sequence redundantly to DBJ
since 1993. For the abundance information of clones with this
sequence in this library and as well as in other 3'-directed
libraries, see 'http://www.imcb.osaka-u.ac.jp/bodymap'. The
sequences of the clones represented by this GS sequences is also
found there.

FEATURES

source

Location/Qualifiers
1..39
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="Human adult (K.Okubo)"
/notes="One or more human adult tissue"
2 c 3 g 18 t

BASE COUNT

ORIGIN

Query Match 1.0%; Score 18.2; DB 14; Length 39;
Best Local Similarity 66.7%; Pred. No. 1.5e+07;
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1762 ACCGTTAACTTAATCTAGCTAGCAATATAAAATC 1800
DB 39 AACTTAAATATATATCATCTTATTAATATAAGATC 1

RESULT 12

AI022013/c

LOCUS

DEFINITION

AI022013 40 bp mRNA linear EST 28-AUG-1998
Ow64c01.x1 Soares senescent fibroblasts NBHSF Homo sapiens cDNA
clone IMAGE:1651584 3' similar to TR:Q35993 Q35993 CYTOCHROME C
OXIDASE III. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 400 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..40
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1651584"
/tissue type="senescent fibroblast"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares senescent fibroblasts_NBHSF"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker V.TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(dT)

FEATURES

source

Location/Qualifiers
1..40
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1651584"
/tissue type="senescent fibroblast"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares senescent fibroblasts_NBHSF"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker V.TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(dT)

primer [5'
TCTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT

ORIGIN

Query Match 1.0%; Score 18.2; DB 9; Length 40;
Best Local Similarity 66.7%; Pred. No. 1.5e+07;
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 824 TGGCGTGTTCGGAAGTGTTCCTTTTCTTTAGGTCGGTT 862
DB 40 TGTGATTTTATTGTGAGTATTTTGTTCCTTTAGTGT 2

RESULT 13

AL849593/c

LOCUS

DEFINITION

AL849593 XGC-egg Silurana tropicalis cDNA clone TBgg061n22 5', mRNA
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 37)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2002
Unpublished
Contact: Taylor R
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE.ID: TEGg061n22.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..37
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGg061n22"
/dev_stage="egg"
/lab_host="Escherichia coli XLI-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5' of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

FEATURES

source

Location/Qualifiers
1..37
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGg061n22"
/dev_stage="egg"
/lab_host="Escherichia coli XLI-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5' of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

BASE COUNT

ORIGIN

Query Match 1.0%; Score 18; DB 9; Length 37;
Best Local Similarity 70.6%; Pred. No. 1.7e+07;
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1378 AGCTCTGTTTATAGCACATTTCTTTTGAAGAAA 1411
DB 34 ASCATTGTTAGATATATCTTCTTTGAAGAAA 1

RESULT 14

AU266702


```

LOCUS       AU266702               39 bp  mRNA  linear  EST 10-MAY-2002
DEFINITION   AU266702 VS Dictyostelium discoideum cDNA clone VSG660 5', mRNA
ACCESSION   AU266702
VERSION     AU266702.1  GI:20525500
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum
ORGANISM    Dictyostelium discoideum
REFERENCE   1 (bases 1 to 39)
AUTHORS     Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
TITLE       Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
JOURNAL     Unpublished
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.

FEATURES             source
BASE COUNT           14 a      0 g      20 t      5 others
ORIGIN
1  .39
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="VSG660"
/sex="mat A"
/dev_stage="vegetative"
/clone_lib="VS"

Query Match      1.0%; Score 18; DB 9; Length 39;
Best Local Similarity 70.0%; Pred. No. 1.7e+07;
Matches 21; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY  1396  TTCTCTTTTGAAGAAATTCAAATAGATT 1425
Db  1  TTTTCTTTTAAAAAATNAAANNTTT 30

RESULT 15
AV853005
LOCUS       AV853005               39 bp  mRNA  linear  EST 08-NOV-2001
DEFINITION   AV853005 Nori Satch unpublished cDNA library, egg Ciona
            intestinalis cDNA clone rcieg16c15 3', mRNA sequence.
ACCESSION   AV853005
VERSION     AV853005.1  GI:16838721
KEYWORDS    EST.
SOURCE      Ciona intestinalis
ORGANISM    Ciona intestinalis
REFERENCE   1 (bases 1 to 39)
AUTHORS     Satch,N., Satou,Y., Kohara,Y. and Shin-i.T.
TITLE       Expressed genes in Ciona intestinalis
JOURNAL     Unpublished
COMMENT     Contact: Nori Satch
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satch@acidian.zool.kyoto-u.ac.jp.

FEATURES             source
BASE COUNT           1. .39
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcieg16c15"

```

```

/tissue_type="whole animal"
/dev_stage="egg"
/clone_lib="Nori Satch unpublished cDNA library, egg"

BASE COUNT           13 a      12 c      3 g      10 t      1 others
ORIGIN
1  108  CGAGCCCTATTTAAACGTCCTCCATTCGAATAATCAT 142
2  2  GCACCCCTATANATAACTTCCCAATGACAACTT 36

Query Match      1.0%; Score 18; DB 9; Length 39;
Best Local Similarity 68.6%; Pred. No. 1.7e+07;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY  108  CGAGCCCTATTTAAACGTCCTCCATTCGAATAATCAT 142
Db  2  GCACCCCTATANATAACTTCCCAATGACAACTT 36

RESULT 16
BZ761743/c
LOCUS       BZ761743               39 bp  DNA  linear  GSS 13-MAR-2003
DEFINITION   SALK 080903.44.85.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_080903.44.85.x, genomic
            survey sequence.
ACCESSION   BZ761743
VERSION     BZ761743.1  GI:28934296
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (chale cress)
ORGANISM    Arabidopsis thaliana
REFERENCE   1 (bases 1 to 39)
AUTHORS     Alonso,J.M., Leishe,T.J., Barajas,P., Chen,H., Cheuk,R., Gadriab,
            C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
            Zimmerman,J. and Ecker,J.R.
TITLE       A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
JOURNAL     Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGNAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.

Class: TDNA tagged
Location/Qualifiers
1. .39
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_080903.44.85.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT           20 a      5 c      2 g      12 t

Query Match      1.0%; Score 18; DB 29; Length 39;
Best Local Similarity 70.6%; Pred. No. 1.7e+07;
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  1395  ATTCTCTTTTGAAGAAATTCAAATAGATTGCA 1428
Db  39  ATTCTCTTTTATGATATTAATAGATGCGTTCA 6

RESULT 17

```

AZ647867/c
 LOCUS
 DEFINITION
 1M0514B09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0514B09 R, genomic survey sequence.
 AZ647867
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 36)
 REFERENCE
 AUTHORS
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 , M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.
 TITLE
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL
 COMMENT
 Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0514 row: B column: 09
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 36.
 Location/Qualifiers
 1..36
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0514B09"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

FEATURES source

1..36
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0514B09"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 12 a 3 c 6 g 15 t
 ORIGIN

Query Match 1.0%; Score 17.8; DB 28; Length 36;
 Best Local Similarity 75.9%; Pred. No. 1.8e+07;
 Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1392 CACATTTCTTTTGAAGAAATTCAAATA 1420
 ||||| |
 DB 30 CACATATAGTTTGAAGAAATGTTTCAATA 2

RESULT 18
 AZ448035/c
 LOCUS
 DEFINITION
 1M0245D02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0245D02 R, genomic survey sequence.
 AZ448035
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 37)
 REFERENCE
 AUTHORS
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 , M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.
 TITLE
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL
 COMMENT
 Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0245 row: D column: 02
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 37.
 Location/Qualifiers
 1..37
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0245D02"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

FEATURES source

1..37
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0245D02"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 12 a 17 c 0 g 8 t
 ORIGIN

Query Match 1.0%; Score 17.8; DB 28; Length 37;
 Best Local Similarity 90.5%; Pred. No. 1.8e+07;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 892*TTTGGGAATGATTTTAAAGG 912
 |
 DB 28 TGTGGGAATGATTTTGAAGG 8

```

RESULT 19
H57900
LOCUS
DEFINITION
yrl4c09.r1 Soares fetal liver spleen INFLS EST 05-OCT-1995
IMAGE:205264 S' similar to gb:M21121 T-CELL SPECIFIC RANTES PROTEIN
PRECUSOR (HUMAN);, mRNA sequence.
H57900
ACCESSION
H57900.1 GI:1010732
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 38)
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished
TITLE
Contact: Wilson RK
JOURNAL
Washington University School of Medicine
COMMENT
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1120
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 1120 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES
Location/Qualifiers
1..38
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GB:3780072"
/db_xref="taxon:9606"
/clone="IMAGE:205264"
/sex="male"
/dev stage="20 week-post conception fetus"
/lab host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' ACTCGAGAGTAATTAAGACTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."
BASE COUNT 3 a 5 c 8 g 21 t 1 others
ORIGIN
Query Match 1.0%; Score 17.8; DB 14; Length 38;
Best Local Similarity 65.8%; Pred. No. 1.8e+07;
Matches 25; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1608 TTTTGTATGCTTTTGTATTATAGCGGCTCTCTCTC 1645
||||| | | | | | | | | | | | | | | |
Db 1 TTTTGTATGCTTTTGTATTATAGCGGCTCTCTCTC 38

RESULT 20
BH864598
LOCUS
DEFINITION
SALK_096314 Arabidopsis thaliana TDNA insertion lines Arabidopsis
thaliana genomic clone SALK_096314, genomic survey sequence.
BH864598
ACCESSION
BH864598.1 GI:22100496
KEYWORDS
GSS
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsais.
REFERENCE
1 (bases 1 to 38)
AUTHORS
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadtrina
b,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
,Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished
TITLE
Contact: Joseph R. Ecker
JOURNAL
Salk Institute Genomic Analysis Laboratory (SIGNAL)
COMMENT
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
FEATURES
Location/Qualifiers
1..38
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_096314"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 12 a 1 c 10 g 15 t
ORIGIN
Query Match 1.0%; Score 17.8; DB 28; Length 38;
Best Local Similarity 67.6%; Pred. No. 1.8e+07;
Matches 25; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1789 AATATAAAATCTCTTTGGGACTTTAGTCCCAAGG 1825
||||| | | | | | | | | | | | | | | |
Db 2 ATTTAGAGATCTTTATGGGATTTAGTGAAAAAGG 38

RESULT 21
AG227568
LOCUS
DEFINITION
Lotus japonicus DNA, clone:LjB141114_f, genomic survey sequence.
ACCESSION
AG227568
VERSION
AG227568.1 GI:26538192
KEYWORDS
GSS.
SOURCE
Lotus japonicus
ORGANISM
Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
REFERENCE
1
AUTHORS
Sato,S., Nakamura,Y. and Tabata,S.
TITLE
Lotus japonicus BAC End sequences
JOURNAL
Published Only in Database (2002)
REFERENCE
2 (bases 1 to 39)
AUTHORS
Sato,S.
TITLE
Direct Submission
JOURNAL
Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,

```

The First Laboratory for Plant Gene Research; 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: ssato@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>,
Tel: 81-438-52-3935(ex.2336), Fax: 81-438-52-3934)

FEATURES

source
Location/Qualifiers

1. .39
/organism="Lotus japonicus"
/mol_type="genomic DNA"
/strain="Miyakojima MG-20"
/db_xref="taxon:34305"
/clone="LjB141114 f"
/clone_lib="genomic BAC library"
/notes="VECTOR:pBelobAC11"
10 a 6 c 5 g 18 t

BASE COUNT

ORIGIN

Query Match 1.0%; Score 17.8; DB 29; Length 39;
Best Local Similarity 67.6%; Pred. No. 1.8e+07;

Matches 25; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1374 GTGCAGTCTTGTATAGACATTTCTTTTGAGAA 1410

Db 3 GTTCACAATTATGATAACACTTTCTTTCTTAGGA 39

RESULT 22

AZ767521

LOCUS AZ767521 40 bp DNA linear GSS 16-FEB-2001
DEFINITION IM0566D18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0566D18 R, genomic survey sequence.

ACCESSION AZ767521

VERSION AZ767521.1 GI:12885699

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 40)

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.,
and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished

JOURNAL

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5506

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0566 row: D column: 18

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 40.

Location/Qualifiers

FEATURES

source

1. .40
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0566D18"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
<http://www.jax.org/resources/documents/dnares/>. The DNA
was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent *E. coli* XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 6 a 5 c 8 g 21 t

ORIGIN

Query Match 1.0%; Score 17.8; DB 28; Length 40;

Best Local Similarity 67.6%; Pred. No. 1.8e+07;

Matches 25; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 822 TTTCGGCTCTTCTGGAAGTGTCTTTCTTTAGTTG 858

Db 2 TCTGCACATAATCTGGAAGTGTCTTTCTTTAGTTG 38

RESULT 23

AU259110/c

LOCUS AU259110 34 bp mRNA linear EST 25-APR-2002
DEFINITION AU259110 3'-directed mouse cDNA library Mus musculus cDNA clone
BED0014559 3', mRNA sequence.

ACCESSION AU259110

VERSION AU259110.1 GI:20325321

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 34)

AUTHORS Kato, K. and Matoba, R.
TITLE Generation of expressed sequence tags from mouse brain
JOURNAL Unpublished

COMMENT Contact: Kikuya Kato

Graduate School of Biological Sciences

Nara Institute of Science and Technology

8916-5 Takayama, Ikoma, Nara 630-0101, Japan

Tel: 81-743-72-5581

Fax: 81-743-72-5589

Email: kkato@bs.aist-nara.ac.jp

URL: <http://love2.aist-nara.ac.jp/BED/index.html>.

FEATURES

source

1. .34
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="BED0014559"
/tissue_type="brain"
/clone_lib="3'-directed mouse cDNA library"
16 a 6 c 3 g 8 t 1 others

BASE COUNT

ORIGIN

Query Match 0.9%; Score 17.6; DB 9; Length 34;

Best Local Similarity 69.7%; Pred. No. 2e+07;

Matches 23; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 706 TTTCATGACTAAATTTTCTTCATGCTGTGGAT 738

Db 34 TTTCATGACTAAATGATGATGCTTCCTGCTAGTAGAT 2

RESULT 24

AG218655

LOCUS

AG218655 34 bp DNA linear GSS 03-SEP-2002

DEFINITION Drosophila melanogaster DNA, clone:NP2234-3-1, flanking P[GawB] transposon insertion, genomic survey sequence.

ACCESSION AG218655

VERSION AG218655.1 GI:22765734

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

REFERENCE 1 Hayashi,S., Ito,K., Sado,Y., Taniguchi,M., Akimoto,A., Takeuchi,H., Aigaki,T., Matsuzaki,F., Nakagoshi,H., Tanimura,T., Ueda,R., Uemura,T., Yoshihara,M. and Goto,S.

TITLE GETDB, a database compiling expression patterns and molecular locations of a collection of Gal4 enhancer traps

JOURNAL Genesis (2002) In press

REFERENCE 2 (bases 1 to 34)

AUTHORS Hayashi,S.

TITLE Direct Submission

JOURNAL Submitted (27-AUG-2002) Shigeo Hayashi, RIKEN Center for Developmental Biology, Laboratory for Morphogenetic Signaling; Chuo-ku, Minatojima-minamimachi 2-2-3, Kobe, Hyogo 650-0047, Japan (E-mail:shayashi@cdb.riken.go.jp, Tel:81-78-301-3184, Fax:81-78-301-3183)

COMMENT This clone was isolated from genomic DNA flanking an insertion of the P element vector P[GawB] of a Drosophila strain.

FEATURES

source

1..34

Location/Qualifiers

organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/strain="NP2234"

/db_xref="taxon:7227"

/chromosome="3"

/map="93C5"

/clone="NP2234-3-1"

/note="flanking P[GawB] transposon insertion"

BASE COUNT 15 a 6 c 3 g 9 t 1 others

ORIGIN

Query Match 0.9%; Score 17.6; DB 29; Length 34;

Best Local Similarity 69.7%; Pred. No. 2e+07;

Matches 23; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 528 CTCGAAGCTTAAATAATCTGTAGAAAAAGGAT 560

||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1 CTCGAAGCTTAAATAATTAGTATAACAGAT 33

RESULT 25

BZ763041

LOCUS BZ763041.1 GI:28935594

DEFINITION SALK_111806.16.85.n Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_111806.16.85.n, genomic survey sequence.

ACCESSION BZ763041

VERSION BZ763041.1 GI:28935594

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 35)

AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrihab ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.

TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL Unpublished

COMMENT Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@alk.edu

This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.

Location/Qualifiers

1..37

organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_139715.21.10.x"

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@alk.edu

This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 3' end of Atg18600.

Class: TDNA tagged.

Location/Qualifiers

1..35

organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_111806.16.85.n"

/note="PCR was performed on Arabidopsis thaliana TDNA insertion lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 12 a 4 c 2 g 17 t

ORIGIN

Query Match 0.9%; Score 17.6; DB 29; Length 35;

Best Local Similarity 71.9%; Pred. No. 2e+07;

Matches 23; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1647 TTAATTGTCATTATTATGACCATAGGAGT 1678

||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 4 TAAATATCCATTATTGTCATATGAGT 35

RESULT 26

BZ768062

LOCUS BZ768062.1 GI:28941108

DEFINITION SALK_139715.21.10.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_139715.21.10.x, genomic survey sequence.

ACCESSION BZ768062

VERSION BZ768062.1 GI:28941108

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 37)

AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrihab ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.

TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL Unpublished

COMMENT Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@alk.edu

This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.

Location/Qualifiers

1..37

organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_139715.21.10.x"

/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html

BASE COUNT 10 a 3 c 9 g 15 t

Query Match 0.9%; Score 17.6; DB 29; Length 37;
Best Local Similarity 71.9%; Pred. No. 2e+07;
Matches 23; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 488 GTTCTTGATTCCGAAGATCCGAGGAGATT 519

DB 1 GTTCTTGCAATTTCTAAGATGTGAGTGAGATT 32

RESULT 27
BH789828
LOCUS
DEFINITION
SALK_046750.38.30.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_046750.38.30.x, genomic
survey sequence.

ACCESSION BH789828.1 GI:19882926

VERSION GSS.

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 40)

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab

,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.

, Zimmerman,J. and Ecker,J.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA.

Class: TDNA tagged.

Location/Qualifiers

1..40

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_046750.38.30.x"

/note="PCR was performed on Arabidopsis thaliana lines

each of which contains one or more TDNA insertion

elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at

the site of insertion. Details of the protocols used can

be found at http://signal.salk.edu/tdna_protocols.html

19 a 2 c 4 g 15 t

Query Match 0.9%; Score 17.6; DB 28; Length 40;

Best Local Similarity 65.0%; Pred. No. 2e+07;

Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1385 TTTATAGACATTTCTTTTGAAGAAATTCAAATGATT 1424

DB 1 TTATTAGAAATTTCTTTTATTTGACAAAAGTAATAGAAATT 40

RESULT 28
BZ663364/c

LOCUS

DEFINITION

SALK_026961.47.05.x Arabidopsis thaliana TDNA insertion lines

Arabidopsis thaliana genomic clone SALK_026961.47.05.x, genomic

survey sequence.

ACCESSION BZ663364.1 GI:28178772

VERSION GSS.

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 40)

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab

,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.

, Zimmerman,J. and Ecker,J.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA.

Class: TDNA tagged.

Location/Qualifiers

1..40

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_026961.47.05.x"

/note="PCR was performed on Arabidopsis thaliana lines

each of which contains one or more TDNA insertion

elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at

the site of insertion. Details of the protocols used can

be found at http://signal.salk.edu/tdna_protocols.html

20 a 7 c 3 g 10 t

Query Match 0.9%; Score 17.6; DB 29; Length 40;

Best Local Similarity 71.9%; Pred. No. 2e+07;

Matches 23; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1466 TTAACAATACCTTTCATCATTTTTCATCGTAT 1497

DB 37 TCAGACAATGATTTCATCATGTTTGTGTTAT 6

RESULT 29

LOCUS

DEFINITION

AA863634

IMAGE:1263648 5' similar to SW:GALE_HUMAN Q14376 UDP-GLUCOSE

4-EPIMERASE ; mRNA sequence.

ACCESSION AA863634.1 GI:2956113

VERSION AA863634

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 28)

AUTHORS

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilton, R. and Waterston, R.

TITLE

The wasNU-HHMI Mouse EST Project

JOURNAL

Unpublished

COMMENT

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:666200

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 23.

FEATURES

source

1. .28
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1263648"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares thymus 2NbMT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 0.9%; Score 17.4; DB 9; Length 28;

Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY

1803 CTTGGGACTTTAGTCCCAAGGCCCC 1829

||||| ||||| ||||| ||||| ||||| |||||

Db 2 CTTGGGTACTTCAATCCCATAGGCGCC 28

RESULT 30

BH866003

LOCUS

DEFINITION SALK_100228 Arabidopsis thaliana DNA insertion lines Arabidopsis thaliana genomic clone SALK_100228, genomic survey sequence.

ACCESSION BH866003

VERSION BH866003.1

KEYWORDS GI:22101901

SOURCE GSS.

ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

REFERENCE 1

AUTHORS

Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadriab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J.R.

A Sequence-indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished

COMMENT

Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.

FEATURES

source

Class: TDNA tagged.

Location/Qualifiers

1. .35

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_100228"

/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 0.9%; Score 17.4; DB 28; Length 35;

Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY

241 AGAGACGATCTCTCTGATCTTTGGAA 267

||||| ||||| ||||| ||||| |||||

Db 7 ACAGACGAGCTCTGTAAGTTGGAA 33

RESULT 31

BH864481

LOCUS

DEFINITION SALK_096115 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_096115, genomic survey sequence.

ACCESSION BH864481

VERSION BH864481.1

KEYWORDS GI:22100379

SOURCE GSS.

ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

REFERENCE 1

AUTHORS

Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadriab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J.R.

A Sequence-indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished

JOURNAL

COMMENT

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA.

Class: TDNA tagged.

Location/Qualifiers

1. .38

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_096115"

```

1 39
/organism="Mus musculus"
/mol_type="Genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0574P02"
/sex="Male"
/lab_host="E. Coli strain XU10-Gold, T1-resistant, F-".
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: pWP42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel

```


AU265359 31 bp mRNA linear EST 10-MAY-2002
AU265359 VS Dictyostelium discoideum cDNA clone VS9523 5', mRNA
sequence.
AU265359

Query Match 0.9%; Score 17.2; DB 28; Length 30;
Best Local Similarity 73.3%; Pred. No. 2.4e+07;
Matches 22; Conservative 0; Mismatches 8; Indels

AU265359 31 bp mRNA linear EST 10-MAY-2002
AU265359 VS Dictyostelium discoideum cDNA clone VS9523 5', mRNA
sequence.
AU265359

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

BASE COUNT      15 a      6 c      8 g      9 t
ORIGIN
Query Match      0.9%; Score 17.2; DB 28; Length 38;
Best Local Similarity 86.4%; Pred. No. 2.5e+07;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1385 TTTATAGCACATTTCTTTTGA 1406
| | | | | | | | | | | | | | | |
Db 38 TGTATAGCACAAATTTTTTTTGA 17

```

RESULT	41
CC0571135/c	
LOCUS	
DEFINITION	CC0571135 39 bp DNA linear GSS 02-APR-2003 SALK_111730.14.10.n Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_111730.14.10.n, genomic survey sequence.
ACCESSION	CC0571135
VERSION	CC057135.1 GI:29476919
KEYWORDS	GSS.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 39)
REFERENCE	Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadri nab A Sequence-Indexed Library of Insertion Mutations in the
AUTHORS	, C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P. Zimmerman,J. and Ecker,J.R.
TITLE	Arabidopsis Genome
COMMENT	Unpublished
	Contact: Joseph R. Ecker

```

Class: tdna tagged.
FEATURES             Location/Qualifiers
     source           1..39
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /strain="Columbia 0"
                        /db_xref="taxon:3702"
                        /clone="SALK_111730.14.10.n"
                        /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                        /note="PCR was performed on Arabidopsis thaliana lines
                        /note="each of which contains one or more TDNA insertion
                        /note="elements. The resultant fragment for each line was
                        /note="directly sequenced to determine the genomic sequence at
                        /note="the site of insertion. Details of the protocols used can
                        /note="be found at http://signal.salk.edu/tdna\_protocols.html"
BASE COUNT          9 a      8 c      12 g      10 t
ORIGIN
Query Match          0.9%; Score 17.2; DB 29; Length 39;
Best Local Similarity 65.8%; Pred. No. 2.5e+07;
Matches 25; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

RESULT 42	AL752023	39 bp	DNA	linear	GSS 17-JUN-2002
LOCUS	AL752023				
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence GK-010D05-014861, genomic survey sequence.				
ACCESSION	AL752023				
VERSION	AL752023.1	GI:21484521			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi				
REFERENCE	1	Strizhov, N., Li, Y., Rosso, M., Viehoever, P., Dekker, K., Suedler, H. and Weisshaar, B.			
AUTHORS	A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines				
TITLE	Unpublished				
JOURNAL	2	Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.			
REFERENCE	A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics				
AUTHORS	Unpublished				
TITLE	3	(bases 1 to 39)			
JOURNAL	Strizhov, N., Li, Y., Rosso, M. and Weisshaar, B.				
REFERENCE	Direct Submission				
AUTHORS	Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany				
TITLE	This sequence is recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by clone F2009. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:				
JOURNAL	http://www.mpiz-koeln.mpg.de/GABI-Kat/.				
COMMENT					

BASE COUNT	removed"			14 t
	7 a	10 c	8 g	
ORIGIN				
Query Match	0.9%; Score 17.2; DB 29; Length 39;			
Best Local Similarity	65.8%; Pred. No. 2.5e+07;			
Matches	25;	Conservative	0;	Mismatches 13; Indels 0; Gaps 0;
Qy	403 AGCTCTACAGTATTTGGATGCGCCACTCACCTTGTGCAG 440			
Db	1 AGCACTGCGGTATCTGGAAGTGTTCCTCTTTCTTCAG 18			

RESULT 43	LOCUS	DEFINITION	40 bp	mRNA	linear	EST 17-FEB-1997
AAL16378						
AAL16378						
mq07c11.r1	Soares thymus 2NDMT	Mus musculus CDNA clone IMAGE:578036				
5'	similar to SW/GATE RAT	pi8645 UDP-GLUCOSE 4-PIPERASE				

sequence.
 AA116378
 VERSION
 AA116378.1 GI:1671391
 EST.
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 40)
 REFERENCE
 AUTHORS
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenger, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE
 The WashU-HMI Mouse EST Project
 JOURNAL
 Unpublished
 COMMENT
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:352684
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28M13 rev2 from AmerSham
 High quality sequence stop: 1.
 Location/Qualifiers
 1..40
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:578036"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares_thymus_2NbMT"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA.
 was primed with a Not I - oligo(dT) primer [5'
 TCTTACCATCTGAGTGGAGCGCGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M.Fatima Bonaldo."
 7 a 15 c 9 g 9 t
 BASE COUNT 7 a 15 c 9 g 9 t
 ORIGIN
 Query Match 0.9%; Score 17.2; DB 9; Length 40;
 Best Local Similarity 73.3%; Pred. No. 2.5e+07;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1800 CTCCTTTGGACTTATAGTCCCAAGGCCCC 1829
 |||||
 Db 8 CTGCTTCGGTACTACATCCCATAGCGCC 37
 |||||
 RESULT 44
 AJ237274/c
 LOCUS
 AJ237274 Emericella nidulans FGSC4 early sexual
 DEFINITION
 cDNA clone esd0690, mRNA sequence.
 ACCESSION
 AJ237274
 VERSION
 AJ237274.1 GI:4191243
 KEYWORDS
 EST.
 SOURCE
 Emericella nidulans (anamorph: Aspergillus nidulans)

ORGANISM Emericella nidulans
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; Emericella.
 REFERENCE
 1 (bases 1 to 40)
 AUTHORS
 Jeong, H., Han, D. M., Jahng, K. and Chae, K.
 TITLE
 The rpl6a Gene for Ribosomal Protein L16A Identified from
 Expressed Sequence Tags is Differentially Expressed during Sexual
 Development of Aspergillus nidulans
 Fungal Genet. Biol. 31 (2), 69-78 (2000)
 JOURNAL
 MEDLINE
 21110325
 PUBMED
 11170736
 COMMENT
 Contact: Chae KS
 Faculty of Biological Sciences
 Chonbuk National University
 Chonju 561-756, Republic of Korea
 POLYA=No.
 FEATURES
 source
 1..40
 /organism="Emericella nidulans"
 /mol_type="mRNA"
 /strain="FGSC4"
 /db_xref="taxon:162425"
 /clone="esd0690"
 /dev_stage="early sexual"
 /clone_lib="Emericella nidulans FGSC4 early sexual"
 BASE COUNT 18 a 8 c 3 g 11 t
 ORIGIN
 Query Match 0.9%; Score 17.2; DB 9; Length 40;
 Best Local Similarity 73.3%; Pred. No. 2.5e+07;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 324 TTTTCATGAGCGTATATTATAGATTGGAG 353
 |||||
 Db 39 TTTTCAGGACTTTATTGTAGATTTCAG 10
 |||||
 RESULT 45
 BI693481/c
 LOCUS
 BI693481 NCI_CGAP_Mam2 Mus musculus cDNA clone
 DEFINITION
 mRNA sequence.
 ACCESSION
 BI693481
 VERSION
 BI693481.1 GI:15656123
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 40)
 REFERENCE
 1 (bases 1 to 40)
 AUTHORS
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11945 row: p column: 15
 High quality sequence stop: 40.
 Location/Qualifiers
 1..40
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:5371646"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"

```

/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam2"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

```

BASE COUNT      11 a      7 c      10 g      12 t
ORIGIN
Query Match      0.9%; Score 17.2; DB 12; Length 40;
Best Local Similarity 73.3%; Pred. No. 2.5e+07;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 475 CGTAGATGCAAAATGTTCTTGTATTTCGAAG 504
Db 30 CGGTGAAGCTGATATTCTTGTATTTCACAG 1

```

Search completed: October 23, 2003, 12:34:16
Job time : 3606 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 08:19:59 ; Search time 6720 Seconds
(without alignments)
11347.551 Million cell updates/sec

Title: US-09-868-987-1
Perfect score: 1864
Sequence: 1 atggacttcgcgattgtgc.....atgacaaattcagataatgc 1864

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 1197658

Minimum DB seq length: 10
Maximum DB seq length: 40

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20.6	1.1	38	6	AX082473 Sequence
C 2	20.4	1.1	33	6	AX004499 Sequence
C 3	20.4	1.1	33	6	AX004508 Sequence
C 4	20.4	1.1	34	6	AX236522 Sequence
C 5	20.4	1.1	40	6	BD179478 Gene expr
C 6	20.4	1.1	40	6	178765 Sequence 21
C 7	20.4	1.1	40	6	192057 Sequence 24
C 8	20.4	1.1	20	6	AR312098 Sequence
C 9	20.4	1.1	20	6	AR314596 Sequence
C 10	20.4	1.1	20	6	AR314597 Sequence
C 11	20.4	1.1	20	6	AR315797 Sequence
C 12	19.8	1.1	40	6	AX394952 Sequence
C 13	19.2	1.0	35	6	BD176659 Promoter.
C 14	19.1	1.0	27	6	AR183406 Sequence
C 15	19.1	1.0	35	6	AX468186 Sequence
C 16	19.1	1.0	40	8	ATH552627 Arabidops
C 17	18.8	1.0	31	6	AR095717 Sequence
C 18	18.8	1.0	31	6	AR145333 Sequence
C 19	18.8	1.0	39	6	AX453607 Sequence
C 20	18.8	1.0	40	6	E27307 DNA polymer
C 21	18.6	1.0	37	6	AX581933 Sequence
C 22	18.6	1.0	37	9	AF505528 Homo sapi
C 23	18.6	1.0	38	6	A22367 oligonucleo
C 24	18.6	1.0	38	6	AR069295 Sequence
C 25	18.6	1.0	38	6	AR076115 Sequence
C 26	18.6	1.0	38	6	I18853 Sequence 34
C 27	18.6	1.0	38	6	I22307 Sequence 34
C 28	18.4	1.0	34	6	AX528937 Sequence
C 29	18.4	1.0	38	6	AR027048 Sequence
C 30	18.4	1.0	40	6	AX515048 Sequence
C 31	18.4	1.0	40	6	AX520873 Sequence
C 32	18.4	1.0	40	6	E27309 DNA polymer
C 33	18.2	1.0	33	6	AX419814 Sequence
C 34	18.2	1.0	34	8	ATH524156 Arabidops
C 35	18.2	1.0	36	6	BD007105 Targeted
C 36	18.2	1.0	36	6	BD007113 Targeted
C 37	18.1	1.0	32	6	AX322217 Sequence
C 38	18.1	1.0	34	6	AX236755 Sequence
C 39	18.1	1.0	38	6	AX580535 Sequence
C 40	17.8	1.0	31	6	AX249610 Sequence
C 41	17.8	1.0	37	6	AR279589 Sequence
C 42	17.8	1.0	37	6	AX002911 Sequence
C 43	17.8	1.0	37	6	AX018237 Sequence
C 44	17.8	1.0	37	6	AX183756 Sequence
C 45	17.8	1.0	38	6	AX228781 Sequence

ALIGNMENTS

RESULT 1
AX082473/c
LOCUS AX082473 38 bp DNA linear PAT 28-FEB-2001
DEFINITION Sequence 11 from Patent WO0111060.
ACCESSION AX082473
VERSION AX082473.1 GI:13184620
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Ravn, P., Madsen, S.M., Vrang, A., Israelsen, H., Johnsen, M.G.,
Bredmose, L., and Arnau, J.
TITLE Method of isolating secretion signals in lactic acid bacteria, and
secretion signals isolated from lactococcus lactis


```
FT          Location/Qualifiers
FEATURES   1..40
source     /organism='Artificial Sequence'
           /mol_type='synthetic construct'
           /db_xref='taxon:32630'
BASE COUNT 7 a 11 c 8 g 14 t
ORIGIN
Query Match 1..11; Score 20.4; DB 6; Length 40;
Best Local Similarity 71.11; Pred. No. 7.2e+06;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 545 TCTGTAGAAAAGGATATACCAAGGCTTTTGGAGCCAT 582
      || ||| ||||| ||||| ||||| ||||| |||||
Db 40 TCATTAGATAAAGGATACATAAAGGCGCGCAGCCAT 3

RESULT 6
LOCUS       I78765               40 bp DNA linear PAT 03-APR-1998
DEFINITION  Sequence 21 from patent US 5693781.
ACCESSION  I78765
VERSION    I78765.1 GI:3014919
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 40)
AUTHORS   Zupancic,T.J. and Yukawa,H.
TITLE     Promoter DNA fragment from coryneform bacteria
JOURNAL   Patent: US 5693781-A 21 02-DEC-1997;
FEATURES   Location/Qualifiers
source     1..40
           /organism='unknown'
BASE COUNT 10 a 10 c 7 g 13 t
ORIGIN

Query Match 1..11; Score 20.4; DB 6; Length 40;
Best Local Similarity 71.11; Pred. No. 7.2e+06;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1038 TTGGATCTTCAGAAAAGATCAAAATCTATTTAGTGAT 1075
      ||||| ||||| ||||| ||||| ||||| |||||
Db 39 TGGATCCGTAGAAAAGATCAAAAGGATCTTCTTGAGAT 2

RESULT 7
LOCUS       I92057/c             40 bp DNA linear PAT 01-DEC-1998
DEFINITION  Sequence 24 from patent US 5726299.
ACCESSION  I92057
VERSION    I92057.1 GI:3936527
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 40)
AUTHORS   Zupancic,T.J. and Yukawa,H.
TITLE     Promoter DNA fragment from coryneform bacteria
JOURNAL   Patent: US 5726299-A 24 10-MAR-1998;
FEATURES   Location/Qualifiers
source     1..40
           /organism='unknown'
BASE COUNT 10 a 10 c 7 g 13 t
ORIGIN

Query Match 1..11; Score 20.4; DB 6; Length 40;
Best Local Similarity 71.11; Pred. No. 7.2e+06;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1038 TTGGATCTTCAGAAAAGATCAAAATCTATTTAGTGAT 1075
      ||||| ||||| ||||| ||||| ||||| |||||
Db 39 TGGATCCGTAGAAAAGATCAAAAGGATCTTCTTGAGAT 2

RESULT 8
LOCUS       AR312098/c           20 bp DNA linear PAT 12-JUN-2003
DEFINITION  Sequence 2635 from patent US 6559294.
ACCESSION  AR312098
VERSION    AR312098.1 GI:31705524
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Griffais,R., Hoiseth,S.K., Zagursky,R.J., Metcalf,B.J., Peek,J.A.,
          Sankaran,B. and Fletcher,L.D.
TITLE     Chlamydia pneumoniae polynucleotides and uses thereof
JOURNAL   Patent: US 6559294-A 2635 06-MAY-2003;
FEATURES   Location/Qualifiers
source     1..20
           /organism='unknown'
BASE COUNT 5 a 6 c 4 g 5 t
ORIGIN

Query Match 1..11; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.5e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 ACGGATGTGCTCAAAATGCG 972
      ||||| ||||| ||||| ||||| |||||
Db 20 ACGGATGTGCTCAAAATGCG 1

RESULT 9
LOCUS       AR314596             20 bp DNA linear PAT 12-JUN-2003
DEFINITION  Sequence 5133 from patent US 6559294.
ACCESSION  AR314596
VERSION    AR314596.1 GI:31708022
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Griffais,R., Hoiseth,S.K., Zagursky,R.J., Metcalf,B.J., Peek,J.A.,
          Sankaran,B. and Fletcher,L.D.
TITLE     Chlamydia pneumoniae polynucleotides and uses thereof
JOURNAL   Patent: US 6559294-A 5133 06-MAY-2003;
FEATURES   Location/Qualifiers
source     1..20
           /organism='unknown'
BASE COUNT 2 a 6 c 4 g 8 t
ORIGIN

Query Match 1..11; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.5e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1709 GCACCACCTCTGTTGTTGTT 1728
      ||||| ||||| ||||| ||||| |||||
Db 1 GCACCACCTCTGTTGTTGTT 20

RESULT 10
LOCUS       AR314597             20 bp DNA linear PAT 12-JUN-2003
DEFINITION  Sequence 5134 from patent US 6559294.
ACCESSION  AR314597
VERSION    AR314597.1 GI:31708023
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
```

REFERENCE 1 (bases 1 to 20)
AUTHORS Griffais, R., Hoiseth, S.K., Zagursky, R.J., Metcalf, B.J., Peek, J.A.,
Sankaran, B., and Fletcher, L.D.
TITLE Chlamydia pneumoniae polynucleotides and uses thereof
JOURNAL Patent: US 6559294-A 5134 06-MAY-2003;
FEATURES Location/Qualifiers
source 1..20
BASE COUNT 5 a 6 c 5 g 4 t
ORIGIN

Query Match 1.1%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.5e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1568 TTCAGCGCGACGTAATGAC 1587
|||||
Db 1 TTCAGCGCGACGTAATGAC 20

RESULT 11
AR315797/c
LOCUS AR315797 20 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 6334 from patent US 6559294.
ACCESSION AR315797
VERSION AR315797.1 GI:31709223
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 20)
AUTHORS Griffais, R., Hoiseth, S.K., Zagursky, R.J., Metcalf, B.J., Peek, J.A.,
Sankaran, B., and Fletcher, L.D.
TITLE Chlamydia pneumoniae polynucleotides and uses thereof
JOURNAL Patent: US 6559294-A 6334 06-MAY-2003;
FEATURES Location/Qualifiers
source 1..20
BASE COUNT 5 a 6 c 4 g 5 t
ORIGIN

Query Match 1.1%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.5e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATTCTGCAAAACCGTGATGG 73
|||||
Db 20 ATTCTGCAAAACCGTGATGG 1

RESULT 12
AX394952
LOCUS AX394952 40 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 89 from Patent WO0218595.
ACCESSION AX394952
VERSION AX394952.1 GI:21066025
KEYWORDS
SOURCE synthetic construct
synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Loomore, S., Wang, J., Bradley, B., Ochs, M., and Yang, Y.P.
TITLE Moraxella polypeptides and corresponding dna fragments and uses thereof
JOURNAL Patent: WO 0218595-A 89 07-MAR-2002;
Aventis Pasteur Limited (CA)
FEATURES Location/Qualifiers
source 1..40
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="5' PCR primer for MCA101924"
BASE COUNT 6 a 2 c 14 g 18 t

ORIGIN

Query Match 1.1%; Score 19.8; DB 6; Length 40;
Best Local Similarity 77.4%; Pred. No. 9.7e+06;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 310 GGCAATGCTTATTGTTTTCATGAGCGTATAT 340
|||||
Db 1 GGCAATTCATATGTTTTCATGAGCGTATAT 31

RESULT 13
BD176659
LOCUS BD176659 35 bp DNA linear PAT 18-MAR-2003
DEFINITION Promoter.
ACCESSION BD176659
VERSION BD176659.1 GI:29122369
KEYWORDS WO 02072819-A/28.
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 35)
AUTHORS Shimojo, T., Takakura, H., Ochiai, K., Asada, K. and Ka, I.
TITLE Promoter
JOURNAL Patent: WO 02072819-A 28 19-SEP-2002;
TAKARA SHUZO CO LTD, TOMOKO SHIMOJO, HIKARU TAKAKURA, KAZUYORI OCHIAI,
KIYOZO ASADA, IKUNOSHIN KATO
COMMENT OS Artificial Sequence
PN WO 02072819-A/28
PF 13-MAR-2002 WO 2002JP002341
PI 14-MAR-2001 JP 01P 072802
PI TOMOKO SHIMOJO, HIKARU TAKAKURA, KAZUYORI OCHIAI, KIYOZO ASADA,
PI IKUNOSHIN KATO
PC C12N15/09, C12N1/19, C12N1/21, C12N5/10, C12P21/02 CC
Description of Artificial Sequence: A sequence of primer AP7R1
CC for
CC amplifying coding region of nitrophenyl phosphatase gene. FH
Key Location/Qualifiers
FT source 1..35
/organism="Artificial Sequence".

FEATURES
source
1..35
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

BASE COUNT 6 a 14 c 10 g 5 t
ORIGIN

Query Match 1.0%; Score 19.2; DB 6; Length 35;
Best Local Similarity 75.0%; Pred. No. 1.3e+07;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 418 GGATCGCCACTACCTTGTCTCAGGACTCGCTG 449
|||||
Db 3 GGACGCGTCACCCCTCTCTGCAGAACTCGCTG 34

RESULT 14
AR183406/c
LOCUS AR183406 27 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 2 from patent US 6340747.
ACCESSION AR183406
VERSION AR183406.1 GI:20226999
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Bazin, H. and Mathis, G.
TITLE Fluorescent conjugates of nucleosides or nucleotides, process for their preparation and their uses
JOURNAL Patent: US 6340747-A 2 22-JAN-2002;

JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gascon Cremieux, 91057 Evry cedex, FRANCE

COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplatne' (<http://www.genoplatne.com> and <http://genoplatne-info.infobiogen.fr>).

FEATURES

source	Location/Qualifiers			
	1..40			
	/organism="Arabidopsis thaliana"			
	/mol_type="genomic DNA"			
	/cultivar="Wassilewskija"			
	/db_xref="taxon:3702"			
	/clone="340D04"			
	/clone_lib="Arabidopsis thaliana T-DNA insertion lines"			
misc_feature	1..40			
	/note="T-DNA flanking sequence"			
	left border"			
BASE COUNT	11 a	7 c	2 g	20 t
ORIGIN				
Query Match	1.0%; Score 19; DB 8; Length 40;			
Best Local Similarity	71.4%; Pred. No. 1.5e+07;			
Matches	25; Conservative	0; Mismatches	10; Indels	0; Gaps 0;
QY	1050	AAAAGATCAAAATCTATTTTAGTGATAAAGCTTTA	1084	
Db	1	AACTTATCAATATCCATCTTATTGATCATTTCTTTA	35	
RESULT 17				
AR095717				
LOCUS	AR095717	31 bp	DNA	linear
DEFINITION	Sequence 12 from patent US 6004804.			
ACCESSION	AR095717			
VERSION	AR095717.1	GI:10023848		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 31)			
AUTHORS	Kumar, R. and Metz, R. A.			
TITLE	Non-chimeric mutational vectors			
JOURNAL	Patent: US 6004804-A 12 21-DEC-1999;			
FEATURES	Location/Qualifiers			
	1..31			
source	/organism="unknown"			
BASE COUNT	6 a	7 c	8 g	10 t
ORIGIN				
Query Match	1.0%; Score 18.8; DB 6; Length 31;			
Best Local Similarity	76.7%; Pred. No. 1.7e+07;			
Matches	23; Conservative	0; Mismatches	7; Indels	0; Gaps 0;
QY	852	TAGGTTTCGCTTGCTCGGGTTTGGAGCCT	881	
Db	2	TAGTTTCAGTCGACTCGAGATTGGATCCT	31	
RESULT 18				
AR145333				
LOCUS	AR145333	31 bp	DNA	linear
DEFINITION	Sequence 12 from patent US 6211351.			
ACCESSION	AR145333			
VERSION	AR145333.1	GI:15107200		
KEYWORDS	Unknown.			
SOURCE	Unknown.			

```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 31)
AUTHORS Kumar, R. and Metz, R. A.
TITLE Chimeric mutational vectors
JOURNAL Patent: US 6211351-A 12 03-APR-2001;
FEATURES
    source
        1..31
        /organism="unknown"
BASE COUNT 6 a 7 c 8 g 10 t
ORIGIN
Query Match 1.0%; Score 18.8; DB 6; Length 31;
Best Local Similarity 76.7%; Pred. No. 1.7e+07;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 852 TAGGTTCCGTTGCTCTCGGTTTGGAGCCT 881
Db 2 TAGGTTTCAGTGCACTCGAGATTGGATCCT 31
RESULT 19
AX53607 AX453607 39 bp DNA linear PAT 06-JUL-2002
LOCUS AX53607
DEFINITION Sequence 7 from Patent WO0231158.
ACCESSION AX453607
VERSION AX453607.1 GI:21712845
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Hanke, P. D.
TITLE Feedback-resistant pyruvate carboxylase gene from Corynebacterium
JOURNAL Patent: WO 0231158-A 7 18-APR-2002;
ARCHER-DANIELS-MIDLAND COMPANY (US)
FEATURES
    source
        1..39
        /organism="Corynebacterium glutamicum"
        /mol_type="genomic DNA"
        /db_xref="taxon:1718"
BASE COUNT 4 a 12 c 10 g 13 t
ORIGIN
Query Match 1.0%; Score 18.8; DB 6; Length 39;
Best Local Similarity 76.7%; Pred. No. 1.6e+07;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1337 GCGGTATGCGCTTTAATTCATGACCTTTG 1366
Db 2 GCGGTATGCGCTTTGTTTCTTTCACCTGATG 31
RESULT 20
E27307/c E27307 40 bp DNA linear PAT 18-JUN-2001
LOCUS E27307
DEFINITION DNA polymerase gene.
ACCESSION E27307
VERSION E27307.1 GI:13026438
KEYWORDS JP 1999151087-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 40)
AUTHORS Yoshizumi, I.
TITLE DNA polymerase gene
JOURNAL Patent: JP 1999151087-A 1 08-JUN-1999;
TAKARA SHUZO CO LTD
COMMENT OS Unidentified
PN JP 1999151087-A/1
PD 08-JUN-1999
FF 19-NOV-1997 JP 1997318665
PR YOSHIZUMI ISHINO
PI C12N15/09, C12N1/21, C12N9/12//(C12N15/09, C12R1.01), (C12N1/21,
PC C12R1.19),
PC (C12N9/12, C12R1.19), C12N15/00, (C12N15/00, C12R1.01) CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..40
/organism="Unidentified".
FEATURES
    source
        1..40
        /organism="Unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT 19 a 2 c 9 g 10 t
ORIGIN
Query Match 1.0%; Score 18.8; DB 6; Length 40;
Best Local Similarity 68.4%; Pred. No. 1.6e+07;
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1646 TTTAATTTGCAATTTATGACCATGAGGATTTCTTCT 683
Db 40 TCTAAGAATTTATTTATTTATCTCCATATGTATCCTCCT 3
RESULT 21
AX581933 AX581933 37 bp mRNA linear PAT 10-JAN-2003
LOCUS AX581933
DEFINITION Sequence 3771 from Patent WO0211674.
ACCESSION AX581933
VERSION AX581933.1 GI:27653743
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
ARTIFICIAL SEQUENCES.
REFERENCE 1
AUTHORS Thompson, J., McSwiggen, J., McKenzie, T., Ayers, D., Szymkowski, D. E.
and Grupe, A.
TITLE Method and reagent for the inhibition of calcium activated chloride
channel-1 (Cic-1)
JOURNAL Patent: WO 0211674-A 3771 14-FEB-2002;
RIBOZYME PHARMACEUTICALS, INC. (US); Syntex (U.S.A.) LLC (US);
Thompson, James (US)
FEATURES
    source
        1..37
        /organism="synthetic construct"
        /mol_type="mRNA"
        /db_xref="taxon:32630"
        /note="Enzymatic Nucleic Acid"
BASE COUNT 10 a 6 c 10 g 11 t
ORIGIN
Query Match 1.0%; Score 18.6; DB 6; Length 37;
Best Local Similarity 72.7%; Pred. No. 1.8e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1812 TTTAGTCCCAAGGCCCTGTGGTATTAAATTT 1844
Db 4 TTCAGGCCGAAAGCGAGTGAGGTCTTAACTTT 36
RESULT 22
AF505528 AF505528 37 bp DNA linear PRI 01-MAY-2003
LOCUS AF505528
DEFINITION Homo sapiens clone Y1.7 T cell receptor gamma CDR3 region sequence.
ACCESSION AF505528
VERSION AF505528.1 GI:30269382
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 37)
AUTHORS Dare,R.
TITLE Direct Submission
JOURNAL Submitted (17-APR-2002) Hematology & Genetic Pathology, Flinders
University of South Australia, Flinders Medical Centre, Flinders
Drive, Bedford Park, South Australia 5042, Australia
FEATURES
source
    Location/Qualifiers
        1..37
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="Y1.7"
misc_feature 1..37
    /cell_type="peripheral blood T cell"
misc_feature 1..12
    /note="T cell receptor gamma CDR3 region"
misc_feature 13..18
    /note="Region: V4"
misc_feature 19..37
    /note="Region: N"
misc_feature 19..37
    /note="Region: J1/2"
BASE COUNT 13 a 6 c 7 g 11 t
ORIGIN
Query Match 1.0%; Score 18.6; DB 9; Length 37;
Best Local Similarity 72.7%; Pred. No. 1.8e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1133 ACCTGGCGTTATGTGGGATTGTTGCAGAAAC 1165
Db 3 ACCTGGGATGATTGTAGAATTATTATGAAGAAC 35

RESULT 23
LOCUS A22367 38 bp DNA linear PAT 04-MAY-1995
DEFINITION oligonucleotide K54 from patent WO9220805.
ACCESSION A22367
VERSION A22367.1 GI:904247
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 38)
AUTHORS St.ang hl.S., Nygren,P.-A., Hansson,M., Uhlen,M. and Nguyen,T.Ngoc.
TITLE Recombinant DNA coding for signal peptide, selective interacting
POLYPEPTIDE AND MEMBRANE ANCHORING SEQUENCE
JOURNAL Patent: WO 9220805-A 4 26-NOV-1992;
FEATURES
source
    Location/Qualifiers
        1..38
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
BASE COUNT 6 a 9 c 9 g 14 t
ORIGIN
Query Match 1.0%; Score 18.6; DB 6; Length 38;
Best Local Similarity 72.7%; Pred. No. 1.8e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 848 CTTTATAGTTGCTTCTCGGGTTTGAGCC 880
Db 6 CTTTATAGTCCCTTTGATTTCCAGCTTGTGTC 38

RESULT 24
LOCUS AR069295/c 38 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 34 from patent US 5891631.
ACCESSION AR069295
VERSION AR069295.1 GI:7220183
KEYWORDS

```

```

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 38)
AUTHORS Goldstein,J.L., Brown,M.S., Briggs,M.R. and Wang,X.
TITLE Methods relating tosterol regulatory element binding proteins
JOURNAL Patent: US 5891631-A 34 06-APR-1999;
FEATURES
source
    Location/Qualifiers
        1..38
            /organism="unknown"
BASE COUNT 7 a 4 c 14 g 13 t
ORIGIN
Query Match 1.0%; Score 18.6; DB 6; Length 38;
Best Local Similarity 72.7%; Pred. No. 1.8e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 121 AAACGTCCCATGAAATAATCATGCCAGTGTCTC 153
Db 35 AATCACCCCACTGTAAATATCACCCCACTGTGTC 3

RESULT 25
LOCUS AR076115 38 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 4 from patent US 5958736.
ACCESSION AR076115
VERSION AR076115.1 GI:10002861
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 38)
AUTHORS St.ang hl.S., Nygren,P.-A., Hansson,M., Uhlen,M. and Nguyen,T.Ngoc.
TITLE Recombinant DNA coding for signal peptide, selective interacting
POLYPEPTIDE AND MEMBRANE ANCHORING SEQUENCE
JOURNAL Patent: US 5958736-A 4 28-SEP-1999;
FEATURES
source
    Location/Qualifiers
        1..38
            /organism="unknown"
BASE COUNT 6 a 9 c 9 g 14 t
ORIGIN
Query Match 1.0%; Score 18.6; DB 6; Length 38;
Best Local Similarity 72.7%; Pred. No. 1.8e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 848 CTTTATAGTTGCTTCTCGGGTTTGAGCC 880
Db 6 CTTTATAGTCCCTTTGATTTCCAGCTTGTGTC 38

RESULT 26
LOCUS I18853/c 38 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 34 from patent US 5498696.
ACCESSION I18853
VERSION I18853.1 GI:1599208
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 38)
AUTHORS Briggs,M.R., Brown,M.S., Goldstein,J.L. and Wang,X.
TITLE Sterol regulatory element binding proteins and their use in
screening assays
JOURNAL Patent: US 5498696-A 34 12-MAR-1996;
FEATURES
source
    Location/Qualifiers
        1..38
            /organism="unknown"
BASE COUNT 7 a 4 c 14 g 13 t
ORIGIN

```

Query Match 1.0%; Score 18.6; DB 6; Length 38;
Best Local Similarity 72.7%; Pred. No. 1.8e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 121 AAAGTCCCATTTGAAAAATCATGCCAGTGTCTC 153
||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 35 AATCACCCCACTGTAAAAATCACCCCACTGTCTC 3

RESULT 27
122307/c
LOCUS 122307 38 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 34 from patent US 5527690.
ACCESSION 122307
VERSION 122307.1 GI:1602661
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 38)
AUTHORS Goldstein, J.L., Brown, M.S., Briggs, M.R. and Wang, X.
TITLE Methods and compositions relating to sterol regulatory element binding proteins
JOURNAL Patent: US 5527690-A 34 18-JUN-1996;
FEATURES Location/Qualifiers
source 1..38
BASE COUNT 7 a 4 c 14 g 13 t
ORIGIN

Query Match 1.0%; Score 18.6; DB 6; Length 38;
Best Local Similarity 72.7%; Pred. No. 1.8e+07;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 121 AAAGTCCCATTTGAAAAATCATGCCAGTGTCTC 153
||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 35 AATCACCCCACTGTAAAAATCACCCCACTGTCTC 3

RESULT 28
AX528937
LOCUS AX528937 34 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 6 from Patent WO02061070.
ACCESSION AX528937
VERSION AX528937.1 GI:25172989
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Smith, H.E.
TITLE Environmentally regulated genes of *Streptococcus suis*
JOURNAL Patent: WO 02061070-A 6 08-AUG-2002;
ID-DeLystad, Instituut voor Dierhouderij en Diergezondheid B.V.
(NL)

FEATURES Location/Qualifiers
source 1..34
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="primer"

BASE COUNT 8 a 5 c 8 g 13 t
ORIGIN

Query Match 1.0%; Score 18.4; DB 6; Length 34;
Best Local Similarity 78.6%; Pred. No. 2e+07;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 306 GCTTGGCAATGCTTATTGTTTTCATGAG 333
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 GCTTGGCAATTCATATGTTTTTTTGAG 33

RESULT 29
AR027048
LOCUS AR027048 38 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 60 from patent US 5856138.
ACCESSION AR027048
VERSION AR027048.1 GI:5937888
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 38)
AUTHORS Fukuda, T.
TITLE Human parathyroid hormone muteins and production thereof
JOURNAL Patent: US 5856138-A 60 05-JAN-1999;
FEATURES Location/Qualifiers
source 1..38
BASE COUNT 12 a 5 c 10 g 11 t
ORIGIN

Query Match 1.0%; Score 18.4; DB 6; Length 38;
Best Local Similarity 69.4%; Pred. No. 2e+07;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 957 ATGTTCCTCAAAATGCGTGCAAAAGTTGTGCATAAAC 993
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 ATGGAGTTCAAATGTTTGCCAAAGTTATGGAGTAAC 37

RESULT 30
AX515048
LOCUS AX515048 40 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 1246 from Patent WO02052044.
ACCESSION AX515048
VERSION AX515048.1 GI:23561891
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Nakamura, Y., Sekine, A., Iida, A. and Saito, S.
TITLE Detection of genetic polymorphisms
JOURNAL Patent: WO 02052044-A 1246 04-JUL-2002;
Riken (JP)

FEATURES Location/Qualifiers
source 1..40
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 11 a 5 c 11 g 13 t
ORIGIN

Query Match 1.0%; Score 18.4; DB 6; Length 40;
Best Local Similarity 78.6%; Pred. No. 2e+07;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 514 GGAATTTTATTGCTCTCAAGTCTTAAA 541
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9 GGAGATTTTTTTTGAGTCAAGACTTAAA 36

RESULT 31
AX520873
LOCUS AX520873 40 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 7071 from Patent WO02052044.
ACCESSION AX520873
VERSION AX520873.1 GI:23571551
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.


```

misc_feature 1..34
/notes="T-DNA flanking sequence
left border"
12 a 5 c 5 g 12 t

BASE COUNT
ORIGIN
Query Match 1.0%; Score 18.2; DB 8; Length 34;
Best Local Similarity 74.2%; Pred. No. 2.2e+07;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 534 GTCTTAAAAAATCTGTAGAAAAGGATATAC 564
|||||
Db 34 GTCTTAATATCTCTAGGACAATAATGTAC 4

RESULT 35
BD007105/c
LOCUS Targeted cytotoxic cells. 36 bp DNA linear PAT 31-JAN-2002
DEFINITION BD007105
ACCESSION BD007105
VERSION BD007105.1 GI:18635476
KEYWORDS JP 2001503978-A/20.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 36)
AUTHORS Chen,S.
TITLE Targeted cytotoxic cells
JOURNAL Patent: JP 2001503978-A 20 27-MAR-2001;
COMMENT WAKE FOLEST UNIVERSITY
OS Unidentified
PN JP 2001503978-A/20
PD 27-MAR-2001
PF 23-OCT-1997 JP 1998519647
PR 23-OCT-1996 US 08/740003
PI SHII CHEN
PC C12N15/09,A61K35/14,A61K38/00,A61P31/12,A61P31/18,
PC A61P35/00,
PC A61P43/00,C12N5/10,C12N15/00,C12N5/00,A61K37/02 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..36
/organism="Unidentified".

BASE COUNT 8 a 10 c 10 g 8 t
ORIGIN
Query Match 1.0%; Score 18.2; DB 6; Length 36;
Best Local Similarity 74.2%; Pred. No. 2.2e+07;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 998 GAAGCTGGTCTTTCTCTAGAGACTTCGGTA 1028
|||||
Db 33 GAAGCTGACGTCGCCTCTAGAGAATTCGGTA 3

RESULT 37
AX322217/c
LOCUS Targeted cytotoxic cells. 32 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 90 from Patent EP1162276.
ACCESSION AX322217
VERSION AX322217.1 GI:18093283
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kramer,M.D., Winter,H. and Reinhartz,J.
TITLE Mrna molecules to be used as indicators of the functional and
activation state of t-lymphocytes
JOURNAL Patent: EP 1162276-A 90 12-DEC-2001;
Lynx Therapeutics GmbH (DE)
FEATURES
source 1..32
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 14 a 6 c 6 g 6 t
ORIGIN
Query Match 1.0%; Score 18; DB 6; Length 32;
Best Local Similarity 80.8%; Pred. No. 2.5e+07;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1592 GCTACAACCTATCATCAGTTTGTAT 1617
|||||
Db 27 GCTAGCAGTCTATCAATTTGTGAT 2

RESULT 38
AX236755/c
LOCUS Targeted cytotoxic cells. 34 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 448 from Patent WO0164922.
ACCESSION AX236755

```


[illegible]

VERSION	AX236755.1	GI:15796341
KEYWORDS	synthetic construct synthetic construct artificial sequences.	
SOURCE	ORGANISM	
REFERENCE	1	
AUTHORS	Arico,M.B., Conanducci,M.C., Galeotti,C.C., Masignani,V.C., Guilianini,M.M. and Pizzi,M.C.	
TITLE	Heterologous expression of neisserial proteins	
JOURNAL	Patent: WO 0164922-A 448 07-SEP-2001; Chiron Spa (IT)	
FEATURES	Location/Qualifiers	
source	1..34	
	/organism="synthetic construct"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32630"	
	/note="Oligonucleotide"	4 t
BASE COUNT	13 a 9 c 8 g 4 t	
ORIGIN		
	Query Match 1.0%; Score 18; DB 6; Length 34;	
	Best Local Similarity 70.6%; Pred. No. 2.5e+07;	
	Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;	
QY	443 CTGCTGGGATTGTTCCTGCTATGGGATGCCG 476	
Db	34 CTTTCTCGATTGTTCTTGCCATATGGGATCCGC 1	
RESULT 39		
LOCUS	AX580535/c	
DEFINITION	Sequence 2373 from Patent W00211674.	
ACCESSION	AX580535	
VERSION	AX580535.1	GI:27649737
KEYWORDS	synthetic construct synthetic construct artificial sequences.	
SOURCE	ORGANISM	
REFERENCE	1	
AUTHORS	Thompson,J., Mcswiggen,J., McKenzie,T., Ayers,D., Szymkowski,D.E. and Grupe,A.	
TITLE	Method and reagent for the inhibition of calcium activated chloride channel-1 (clca-1)	
JOURNAL	Patent: WO 0211674-A 2373 14-FEB-2002; RIBOZYME PHARMACEUTICALS, INC. (US) ; Syntex (U.S.A.) LLC (US) ; Thompson, James (US)	
FEATURES	Location/Qualifiers	
source	1..38	
	/organism="synthetic construct"	
	/mol_type="mRNA"	
	/db_xref="taxon:32630"	
	/note="Enzymatic Nucleic Acid"	12 t
BASE COUNT	8 a 8 c 10 g 12 t	
ORIGIN		
	Query Match 1.0%; Score 18; DB 6; Length 38;	
	Best Local Similarity 70.6%; Pred. No. 2.4e+07;	
	Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;	
QY	1808 GGACTTAGTCCAAAGGCCCTGTGGTATTA 1841	
Db	35 GGACTTTCGCCTAACGGCCTCATCAGTATGAA 2	
RESULT 40		
LOCUS	AX249610/c	
DEFINITION	Sequence 1689 from Patent W00166800.	
ACCESSION	AX249610	
VERSION	AX249610.1	GI:15864233
KEYWORDS	Homo sapiens (human)	
SOURCE	ORGANISM	
REFERENCE	1	
AUTHORS	Harmsen,M.M. and Frenken,L.G.	
TITLE	New products comprising inactivated yeasts or moulds provided with active antibodies	
JOURNAL	Patent: EP 0954978-A 11 10-NOV-1999; UNILEVER PLC (GB); UNILEVER NV (NL) Location/Qualifiers	
FEATURES		
	Query Match 1.0%; Score 17.8; DB 6; Length 37;	
	Best Local Similarity 75.9%; Pred. No. 2.7e+07;	
	Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	877 AGCCTTGAATCCGTTTTGGGAATCGATT 905	
Db	8 AGCCTTGGATTCTGTGTAGGATTGGGTT 36	
RESULT 42		
LOCUS	AX002911	
DEFINITION	Sequence 11 from Patent EP0954978.	
ACCESSION	AX002911	
VERSION	AX002911.1	GI:9926827
KEYWORDS	synthetic construct synthetic construct artificial sequences.	
SOURCE	ORGANISM	
REFERENCE	1	
AUTHORS	Harmsen,M.M. and Frenken,L.G.	
TITLE	New products comprising inactivated yeasts or moulds provided with active antibodies	
JOURNAL	Patent: EP 0954978-A 11 10-NOV-1999; UNILEVER PLC (GB); UNILEVER NV (NL) Location/Qualifiers	
FEATURES		
	Query Match 1.0%; Score 17.8; DB 6; Length 37;	
	Best Local Similarity 75.9%; Pred. No. 2.7e+07;	
	Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	877 AGCCTTGAATCCGTTTTGGGAATCGATT 905	
Db	8 AGCCTTGGATTCTGTGTAGGATTGGGTT 36	
ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1	
AUTHORS	Cargill,M., Ireland,J.S. and Lander,E.S.	
TITLE	Human single nucleotide polymorphisms	
JOURNAL	Patent: WO 0166800-A 1889 13-SEP-2001; WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)	
FEATURES	Location/Qualifiers	
source	1..31	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
BASE COUNT	11 a 7 c 3 g 9 t 1 others	
ORIGIN		
	Query Match 1.0%; Score 17.8; DB 6; Length 31;	
	Best Local Similarity 75.9%; Pred. No. 2.7e+07;	
	Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	1400 TTTTCAAGAAAATTCAAATAGATTGCA 1428	
Db	29 TTTTGAAGACAATCAGGTTTATGTCGA 1	
RESULT 41		
LOCUS	AR279589	
DEFINITION	Sequence 11 from patent US 6517829.	
ACCESSION	AR279589	
VERSION	AR279589.1	GI:29714479
KEYWORDS	Unknown.	
SOURCE	ORGANISM	
REFERENCE	1 (bases 1 to 37)	
AUTHORS	Frenken,L.G.J., Harmsen,M.M., van der Linden,R.H.J. and Verrips,C.T.	
TITLE	Products comprising inactivated yeasts or moulds provided with active antibodies	
JOURNAL	Patent: US 6517829-A 11 11-FEB-2003;	
FEATURES	Location/Qualifiers	
source	1..37	
	/organism="unknown"	
BASE COUNT	4 a 6 c 14 g 13 t	
ORIGIN		
	Query Match 1.0%; Score 17.8; DB 6; Length 37;	
	Best Local Similarity 75.9%; Pred. No. 2.7e+07;	
	Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	877 AGCCTTGAATCCGTTTTGGGAATCGATT 905	
Db	8 AGCCTTGGATTCTGTGT	

```
source
1. .37
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="primer_bind"
4 a 6 c 14 g 13 t
BASE COUNT
ORIGIN
Query Match 1.0%; Score 17.8; DB 6; Length 37;
Best Local Similarity 75.9%; Pred. No. 2.7e+07;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 877 AGCCTGGAATTCGTTTGGGAATGGATT 905
|||||
DB 8 AGCCTTGGATTCTGTTGTAGGATTGGGTT 36
|||||

RESULT 43
AX018237
LOCUS AX018237 37 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 11 from Patent WO9946300.
ACCESSION AX018237
VERSION AX018237.1 GI:10042576
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Harmsen,M.M., Convents,D., Frenken,L.G., Van Der Linden,R.H. and
Verrrips,T.C.
TITLE Products comprising inactivated yeasts or moulds and active
vhh-type antibodies
JOURNAL UNILEVER PLC (GB); HARMSSEN MICHAEL MARIE (NL); UNILEVER NV (NL);
CONVENTS DANIEL (NL); FRENKEN LEON GERARDUS JOSEPH (NL); LINDEN
RICHARD HENRICUS JACOBUS (NL); VERRIPS THEODORUS CORNELIS (NL)
FEATURES
source
1. .37
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="primer_bind"
4 a 6 c 14 g 13 t
BASE COUNT
ORIGIN
Query Match 1.0%; Score 17.8; DB 6; Length 37;
Best Local Similarity 75.9%; Pred. No. 2.7e+07;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 877 AGCCTGGAATTCGTTTGGGAATGGATT 905
|||||
DB 8 AGCCTTGGATTCTGTTGTAGGATTGGGTT 36
|||||

RESULT 44
AX183756/c
LOCUS AX183756 37 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1509 from Patent WO0142511.
ACCESSION AX183756
VERSION AX183756.1 GI:15135081
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Daly,M., Hudson,T.J., Lander,E.S., Rioux,J. and Siminovitch,K.
TITLE Ibd-related polymorphisms
JOURNAL Patent: WO 0142511-A 1509 14-JUN-2001;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Ellipsis
Biotherapeutics Corporation (CA)
FEATURES
source
1. .37
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="primer_bind"
4 a 6 c 14 g 13 t
BASE COUNT
ORIGIN
Query Match 1.0%; Score 17.8; DB 6; Length 37;
Best Local Similarity 75.9%; Pred. No. 2.7e+07;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 877 AGCCTGGAATTCGTTTGGGAATGGATT 905
|||||
DB 8 AGCCTTGGATTCTGTTGTAGGATTGGGTT 36
|||||

RESULT 45
AX228781/c
LOCUS AX228781 38 bp mRNA linear PAT 10-SEP-2001
DEFINITION Sequence 2153 from Patent WO0157206.
ACCESSION AX228781
VERSION AX228781.1 GI:15557922
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Fattaey,A.R., Jarvis,T., Mcswiggen,J., Boober,R.N. and Holman,P.S.
TITLE Method and reagent for the inhibition of checkpoint kinase-1 (chk
1) enzyme
JOURNAL Patent: WO 0157206-A 2153 09-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Fattaey, Ali R. (US)
FEATURES
source
1. .38
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/notes="taxon:32630"
16 a 8 c 10 g 4 t
BASE COUNT
ORIGIN
Query Match 1.0%; Score 17.8; DB 6; Length 38;
Best Local Similarity 67.6%; Pred. No. 2.7e+07;
Matches 25; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 583 TTTTGATTCTAACTTGACTACAGTATTGGCCCTCAGCA 619
|||||
DB 38 TTTTGTAGACCTTGACTCGCCCTTCGGCTTCGGCA 2
|||||

Search completed: October 23, 2003, 11:34:00
Job time : 6723 secs
```


DR WPI; 2001-191547/19.

XX Constructing a transposon derivative to identify DNA sequence encoding

PT signal peptide in lactic acid bacteria, involves removing stop codons

PT in frame with secretion reporter molecule from DNA comprising

PT transposon -

XX Example 1; Page 14; 62pp; English.

XX The present invention describes a method for constructing a transposon

CC derivative for identifying DNA (I) encoding a signal peptide (secretion

CC signal, SP) in a lactic acid bacterium (e.g. *Lactococcus lactis*). The

CC method comprises selecting a transposon (II), including a promoterless

CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between

CC its left and right termini (LR) and (RR), deleting a region between LR

CC and PPR gene to obtain modified DNA that retains its transposability and

CC its RBS. The present invention also describes: (i) a transposon

CC derivative (III) for the identification of (I) in a lactic acid

CC bacterium, comprising (II) without stop codons in the region upstream of

CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule

CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)

CC comprising at least a part of (III) and (I) that is functional in a

CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP

CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of

CC the signal peptides having retained signal peptide functionality; (4) a

CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant

CC bacterium (VI) comprising (I). (III) is useful for identifying and

CC isolating (I) from a source lactic acid bacteria, by transforming the

CC bacteria with (III), and selecting from the transformed bacteria, cells

CC in which the promoterless promoter reporter gene is expressed and the

CC gene product of the DNA sequence coding for a secretion reporter molecule

CC is secreted. (VI) is useful for the production of a desired gene product.

CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used

CC in the exemplification of the present invention.

XX

SQ Sequence 38 BP; 14 A; 8 C; 7 G; 9 T; 0 other;

Query Match 1.1%; Score 20.6; DB 22; Length 38;

Best Local Similarity 85.2%; Pred. No. 8.5e+04;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1475 ACATTGATCATTTTTGATCGTATTCGT 1501

||||| ||||| ||||| ||||| |||||

Db 35 ACITTTATATTTTATAGATCGAATTCGT 9

RESULT 2

AAx91515/c

ID AAX91515 standard; DNA; 33 BP.

XX

AC AAX91515;

XX

DT 28-SEP-1999 (first entry)

XX

DE Bovine interferon-alpha gene amplifying primer aIFNsrev.

XX

KW Bovine herpes virus-2; BHV-2; cytokine; immune response; immunogen; IL;

KW interleukin; colony stimulating factor; CSF; interferon; IFN; TNF;

KW tumour necrosis factor; recombinant virus; vaccine; mucosal disease;

KW mastitis; breast cancer; stress-induced disease; PCR primer; ss.

XX

OS Synthetic.

OS Bos sp.

XX

XX WO9916892-A1.

PN

XX

PD 08-APR-1999.

XX

PF 29-SEP-1998; 98WO-GB02927.

XX

PR 29-SEP-1997; 97GB-0020633.

XX

XX (UYBR-) UNIV BRISTOL.

PA

XX Bradley AJ, Duffas WPH;

XX WPI; 1999-255101/21.

DR

XX New bovine herpes virus-2 vectors

XX

XX Example 2; Page 34; 130pp; English.

XX The invention provides bovine herpes virus-2 (BHV-2) based vectors that

CC comprise at least one cytokine-encoding DNA sequence. The expression of

CC cytokines in mammals can up-regulate immune responses to the immunogens.

CC The cytokine is selected from interleukins (IL), colony stimulating

CC factors (CSF), interferons (IFN) and tumour necrosis factors (TNF). The

CC BHV-2 based vector or recombinant virus can be used as vaccines. They can

CC be used for preventing or treating a mucosal disease in a subject, e.g.

CC ruminants, cattle, sheep, goats, dogs, cats, horses, or man, particularly

CC mastitis in cows or breast cancers in humans. They can also be used for

CC preventing or treating a stress-induced disease.

XX

SQ Sequence 33 BP; 10 A; 8 C; 9 G; 6 T; 0 other;

Query Match 1.1%; Score 20.4; DB 20; Length 33;

Best Local Similarity 80.0%; Pred. No. 9.1e+04;

Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 862 TGCTCTCGGTTTGAGCTGCGAATTCGT 891

||||| ||||| ||||| ||||| |||||

Db 33 TGCTCTCGGTTTGACCTGCGAATTCAGT 4

RESULT 3

AAx43955/c

ID AAS43955 standard; DNA; 34 BP.

XX

AC AAS43955;

XX

DT 18-DEC-2001 (first entry)

XX

DE Neisseria meningitidis B MC58 genomic DNA sequence PCR primer #50.

XX

KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ss;

KW Neisserial protein; PCR primer.

XX

OS Neisseria meningitidis.

OS Synthetic.

XX

PN WO200164922-A2.

XX

PD 07-SEP-2001.

XX

PF 28-FEB-2001; 2001WO-IB00452.

XX

PR 28-FEB-2000; 2000GB-0004695.

PR 13-NOV-2000; 2000GB-0027675.

XX

PA (CHIR-) CHIRON SPA.

XX

PI Arico MB, Comanducci M, Galeotti C, Masignani V, Guilianini MM;

PI Pizza M;

XX

DR WPI; 2001-582163/65.

XX

PT Producing heterologous proteins from Neisseria meningitidis and N.

PT gonorrhoeae -

XX

PS Disclosure; Page 82; 119pp; English.

XX The invention relates to methods for the heterologous expression of

CC Neisserial proteins from Neisseria meningitidis and Neisseria

CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the

CC leader peptide, and may be replaced by a domain from a different protein

CC to make a fusion protein, in order to enhance heterologous expression of

Query Match 1.18; Score 20.4; DB 16; Length 40;

ABZ25344/C

AB225344/C
ID AB225344 standard; DNA: 40 BP.

XX PS Page 1777; Disclosure; 1912pp; English.

XX CC AAX91991-X97517 represent PCR primers used to amplify open reading
 CC frames and other nucleic acid sequences from the genome of
 CC Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
 CC disease such as pneumonia and bronchitis and is thought to be a
 CC contributing factor in heart disease, sarcoidosis, sinusitis, purulent
 CC otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
 CC by the open reading frames of the C. pneumoniae genome (see AAY34584-
 CC AAY35879) can be used in immunogenic compositions as vaccines. Vectors
 CC containing C. pneumoniae nucleotides sequences can also be used as
 CC immunogenic compositions, especially where the vector directs the
 CC expression of a neutralising epitope of C. pneumoniae.
 XX CC Sequence 20 BP; 2 A; 6 C; 4 G; 8 T; 0 other;

Query Match 1.1%; Score 20; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.4e+04;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1709 GCACCACCTCTGTGTGTT 1728
 Db 1 GCACCACCTCTGTGTGTT 20

RESULT 9
 AAX95808
 ID AAX95808 standard; DNA; 20 BP.
 XX AC AAX95808;
 XX DT 13-SEP-1999 (first entry)
 XX DE PCR primer used to amplify an ORF of Chlamydia pneumoniae.
 XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 XX KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 XX KW vaccine; neutralising epitope; PCR primer; ss.

XX OS Synthetic.
 OS Chlamydia pneumoniae.

XX PN WO9927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX DR WPI; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae

XX PS Page 1777; Disclosure; 1912pp; English.

XX CC AAX91991-X97517 represent PCR primers used to amplify open reading
 CC frames and other nucleic acid sequences from the genome of
 CC Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
 CC disease such as pneumonia and bronchitis and is thought to be a
 CC contributing factor in heart disease, sarcoidosis, sinusitis, purulent
 CC otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
 CC by the open reading frames of the C. pneumoniae genome (see AAY34584-
 CC AAY35879) can be used in immunogenic compositions as vaccines. Vectors
 CC containing C. pneumoniae nucleotides sequences can also be used as
 CC immunogenic compositions, especially where the vector directs the
 CC expression of a neutralising epitope of C. pneumoniae.

XX SQ Sequence 20 BP; 5 A; 6 C; 5 G; 4 T; 0 other;

Query Match 1.1%; Score 20; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.4e+04;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1568 TTCAGCCGCGGTAATGAC 1587
 Db 1 TTCAGCCGCGGTAATGAC 20

RESULT 10
 AAX93334/c
 ID AAX93334 standard; DNA; 20 BP.
 XX AC AAX93334;
 XX DT 13-SEP-1999 (first entry)
 XX DE PCR primer used to amplify an ORF of Chlamydia pneumoniae.
 XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 XX KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 XX KW vaccine; neutralising epitope; PCR primer; ss.

XX OS Synthetic.
 OS Chlamydia pneumoniae.

XX PN WO9927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX DR WPI; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae

XX PS Page 1581; Disclosure; 1912pp; English.

XX CC AAX91991-X97517 represent PCR primers used to amplify open reading
 CC frames and other nucleic acid sequences from the genome of
 CC Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
 CC disease such as pneumonia and bronchitis and is thought to be a
 CC contributing factor in heart disease, sarcoidosis, sinusitis, purulent
 CC otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
 CC by the open reading frames of the C. pneumoniae genome (see AAY34584-
 CC AAY35879) can be used in immunogenic compositions as vaccines. Vectors
 CC containing C. pneumoniae nucleotides sequences can also be used as
 CC immunogenic compositions, especially where the vector directs the
 CC expression of a neutralising epitope of C. pneumoniae.

XX SQ Sequence 20 BP; 5 A; 6 C; 4 G; 5 T; 0 other;

Query Match 1.1%; Score 20; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.4e+04;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 AGCGATGTTGCTCAATGCG 972
 Db 20 AGCGATGTTGCTCAATGCG 1

RESULT 11
 AAL46489

KW PCR primer; ss.
 XX Synthetic.
 OS Staphylococcus aureus.
 XX WO9945123-A1.
 PN 10-SEP-1999.
 XX 02-MAR-1999; 99WO-US04512.
 XX 02-MAR-1998; 98US-0076525.
 PR (ABBO) ABBOTT LAB.
 XX Aharonowitz Y, Borovok I, Cohen G, Uziel O, Katz L;
 DR WPI; 1999-551044/46.
 XX A new thioredoxin reductase from Staphylococcus aureus -
 PT Example 3; Page 32; 59pp; English.
 PS The present invention describes Staphylococcus thioredoxin reductase
 CC (TrxB). The present sequence represent a PCR primer for S. aureus TrxB.
 CC TrxB inhibitors can be used as antimicrobials to treat a Staphylococcus,
 CC particularly S. aureus, infection. TrxB inhibitors are antimicrobials to
 CC which, unlike most of those in the prior art, Staphylococcus has not yet
 CC developed a resistance.
 XX Sequence 32 BP; 11 A; 3 C; 7 G; 11 T; 0 other;
 SQ Query Match 1.0%; Score 19.4; DB 20; Length 32;
 Best Local Similarity 79.3%; Pred. No. 1.7e+05;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1055 ATCAAAATCTATTTAGTGATTAAGCTTT 1083
 DB 32 ATCAAAATCTATTTAGTGATTAAGCTTT 4
 RESULT 14
 ABQ81130
 ID ABQ81130 standard; DNA; 32 BP.
 XX AC ABQ81130;
 XX 25-NOV-2002 (first entry)
 DE Human G-protein coupled receptor GPR54 PCR primer HGPR54.R8.
 XX GPR54; G-protein coupled receptor; receptor; human; gene therapy;
 KW antidiabetic; analgesic, vasotropic; antimigraine; antidepressant;
 KW nootropic; neuroprotective; tranquilizer; PCR; primer; ss.
 XX Homo sapiens.
 OS WO200259344-A2.
 PN 01-AUG-2002.
 XX 14-DEC-2001; 2001WO-US48333.
 XX 18-DEC-2000; 2000US-256299P.
 PR (MERI) MERCK & CO INC.
 PA Liu Q, Clements M, McDonald TP;
 PI WPI; 2002-666905/71.
 XX New isolated GPR54 polynucleotides and polypeptides, useful for
 PT preventing and/or treating disorders associated with an excess or

PT deficiency of GPR54 protein, such as diabetes, pain, anxiety,
 XX depression and Alzheimer's disease -
 PS Example 1; Page 77; 125pp; English.
 XX The present sequence is that of PCR primer HGPR54.R8, which was
 CC used with primer HGPR54.F6 (see ABQ81129) in a second-round PCR
 CC for the amplification of human G-protein coupled receptor GPR54
 CC cDNA. The template was a first-round PCR (see ABQ81127-28)
 CC product. A clone containing a full-length GPR54 coding sequence
 CC (see ABQ81125) was obtained. The invention provides human and
 CC mouse GPR54 polypeptides and polynucleotides, and methods of
 CC inhibiting or activating these polypeptides and polynucleotides for
 CC use in preventing and/or treating abnormal conditions associated
 CC with excess or insufficient human GPR54 activity, such as eating
 CC disorders, diabetes, pain, migraine, anxiety, depression, ischaemia,
 CC Alzheimer's disease, and reproductive and sleep disorders.
 XX Sequence 32 BP; 12 A; 7 C; 5 G; 8 T; 0 other;
 SQ Query Match 1.0%; Score 19.4; DB 24; Length 32;
 Best Local Similarity 79.3%; Pred. No. 1.7e+05;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1038 TTGGATCTTCAGAAAGATCAAAATCTAT 1066
 DB 4 TTGGATCTTCAGAGACCAAAATATTT 32
 RESULT 15
 AAH75604/c
 ID AAH75604 standard; DNA; 33 BP.
 XX AC AAH75604;
 XX 26-OCT-2001 (first entry)
 DT Human withering related protein 22 PCR primer 5.
 DE Human; cell withering related protein 22; malignant tumour; nosohaemia;
 KW human immunodeficiency virus; HIV; infection; immunological disease;
 KW inflammation; PCR primer; ss.
 XX Homo sapiens.
 OS CN1297906-A.
 PN 06-JUN-2001.
 XX 24-NOV-1999; 99CN-0124099.
 PF 24-NOV-1999; 99CN-0124099.
 PR (SHAN-) SHANGHAI BORONG GEGENE DEV CO LTD.
 XX Mao Y, Xie Y;
 PI WPI; 2001-489652/54.
 DR Human cell withering related protein 22 and encoding polynucleotides,
 XX useful for treating tumour, human immunodeficiency virus infection and
 PT inflammation -
 PS Example 5; Page 18 Disclosure; 27pp; Chinese.
 XX The invention relates to the human cell withering related protein 22. The
 CC polypeptide is useful for treating various diseases, such as malignant
 CC tumour, nosohaemia, human immunodeficiency virus (HIV) infection,
 CC immunological diseases and inflammations. The present sequence is a that
 CC of the human cell withering related protein 22 PCR primer.
 CC Note: The present sequence given as SEQ ID NO 5 in the examples differs
 CC from that given as SEQ ID NO 5 in the sequence listing (AAH75602).
 XX

SQ Sequence 33 BP; 14 A; 6 C; 7 G; 6 T; 0 other;
 Query Match 1.0%; Score 19.4; DB 22; Length 33;
 Best Local Similarity 79.3%; Pred. No. 1.7e+05;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 614 TCAGCATCTCTTTCTTCCTAGATACAGG 642
 DB 29 TCAGCTTTTCTTTCTTCATGGATCCGG 1
 RESULT 16
 ABQ82518
 ID ABQ82518 standard; DNA; 35 BP.
 XX AC ABQ82518;
 XX DT 18-DEC-2002 (first entry)
 XX DE Nitrophenyl phosphatase gene PCR primer AP7R1 SEQ ID NO:28.
 XX KW Promoter: stationary phase-specific promoter; Gram-negative bacteria;
 KW expression; protein synthesis; biotechnological; nitrophenyl phosphatase;
 KW PCR primer; ss.
 XX OS Aeropyrum pernix.
 OS Synthetic.
 XX PN W0200272819-A1.
 XX PD 19-SEP-2002.
 XX 13-MAR-2002; 2002WO-JP02341.
 PF 14-MAR-2001; 2001JP-0072802.
 PR (TAKA-) TAKARA BIO INC.
 XX Shimajo T, Takakura H, Ochiai K, Asada K, Kato I;
 WPI; 2002-723352/78.
 DR Promoters for expression of target gene products, applicable in protein
 PT synthesis by biotechnological methods -
 XX Example 5; Page 45; 54pp; Japanese.
 PS The present invention describes an isolated DNA comprising: (a) an
 CC isolated DNA containing a DNA with a base sequence of any of the defined
 CC sequences (I)-(VI) given in ABQ82491 to ABQ82496, or their fragments, and
 CC showing a stationary phase-specific promoter activity in Gram-negative
 CC bacteria; and (b) an isolated DNA hybridisable with the DNA in (a) under
 CC stringent conditions. Also described: (i) a recombinant DNA containing
 CC the DNA and a foreign gene in which the DNA is so placed as to enable
 CC expression of the foreign gene; (iii) an expression vector for expressing
 CC a gene that contains the DNA; (iii) a transformant cell that can sustain
 CC the recombinant DNA, or expression vector; (iv) a process for producing
 CC a protein by culturing the transformant cells and collecting the product
 CC from the cultured material; and (v) a kit for the protein production
 CC containing the DNA, or the vector for expressing the gene. The promoters
 CC are applicable in protein synthesis by biotechnological methods. With
 CC these promoters, gene products can be conveniently obtained with high
 CC expression, at low cost. The present invention represents a PCR primer
 CC for nitrophenyl phosphatase, which is used in an example from the present
 CC invention.
 XX
 SQ Sequence 35 BP; 6 A; 14 C; 10 G; 5 T; 0 other;
 Query Match 1.0%; Score 19.2; DB 24; Length 35;
 Best Local Similarity 75.0%; Pred. No. 2e+05;
 Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 418 GGATGGCCCACTCACCTTGTGAGGACTCGGTG 449

XX AAX56490;
XX
XX 27-JUL-1999 (first entry)
XX
XX Locked nucleoside analogue oligomer ODN#10.
XX
XX Locked nucleoside analogue; LNA; bicyclic; tricyclic; diagnosis;
KW PCR application; strand displacement oligomer; polymerase; substrate;
KW nucleotide based drug; diagnostic probe; antisense therapy;
KW antiviral; antitumor; ss.
XX
XX Synthetic.
XX
XX WO9914226-A2.
XX
XX 25-MAR-1999.
XX
XX 14-SEP-1998; 98WO-DK00393.
XX
XX 28-JUL-1998; 98DK-0000982.
XX
XX 12-SEP-1997; 97DK-0001054.
XX
XX 19-DEC-1997; 97DK-0001492.
XX
XX 16-JAN-1998; 98DK-0000061.
XX
XX 03-MAR-1998; 98DK-0000286.
XX
XX 29-APR-1998; 98DK-0000585.
XX
XX 05-JUN-1998; 98US-0088309.
XX
XX 08-JUN-1998; 98DK-0000750.
XX
XX (EXIQ-) EXIQON AS.
XX
XX Nielsen P, Wengel J;
XX
XX WPI; 1999-337376/28.
XX
XX New oligonucleotides containing polycyclic, locked nucleoside
PT analogues, useful e.g. as diagnostic probes or in antisense therapy
PT
XX
XX Example 154; Page 175; 269pp; English.
XX
XX The present invention describes novel modified oligonucleotides (I)
CC containing at least one locked nucleoside analog (LNA). Monomeric LNA's
CC (II) are also described. (I) are used: (i) to bind to target sequences
CC in double-stranded DNA or RNA (by strand displacement or triplex
CC formation); (ii) as ribozymes; (iii) as therapeutic antisense, antigene
CC or gene activating agents, specifically for recruitment of RNase H; (iv)
CC diagnostically for isolation, purification, detection, identification,
CC quantitation or capture of (synthetic) nucleic acid, e.g. as probes or
CC primers; (v) as aptamers for therapy, diagnosis, RNA-mediated catalytic
CC processes and for specific binding to antibodies, drugs etc., including
CC resolution of enantiomers; (vi) for labeling, then separating, cells;
CC and (vii) to hybridize to non-coding RNA. LNA are used in synthesis of
CC (I); as therapeutic and diagnostic agents; to equalize the melting point
CC of unmodified reference oligonucleotides and as enzyme substrates.
CC Typical therapeutic applications are as antiviral and antitumor agents.
CC (I) have increased specificity and/or affinity, i.e. higher melting
CC point (Tm), for complementary RNA or DNA than oligomers not containing
CC LNA, and are more resistant to nuclease. The present sequence represents
CC an oligomer used in an example from the present invention.
XX
XX Sequence 35 BP; 5 A; 6 C; 10 G; 14 T; 0 other;
SQ
Query Match 1.0%; Score 19; DB 20; Length 35;
Best Local Similarity 71.4%; Pred. No. 2.2e+05;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1582 AATGACAACAGCTACAACTCTATCAGTTTGTGTTAA 1616
DB 35 AAGTACAACAGCGCAAGTCCAGCACCCTGCTTAA 1
RESULT 19
ABN81648/c

ID ABN81648 standard; DNA; 35 BP.
XX
XX AC
XX ABN81648;
XX
XX 29-AUG-2002 (first entry)
XX
XX Human RGS9 related PCR primer SEQ ID NO 25.
DE
XX RGS; regulator of G-proteins; pheromone; luciferase; Renilla; Photinus;
KW GFP; high through-put screening; human; RGS9; PCR; primer; ss.
KW
XX Synthetic.
OS
XX WO200250104-A2.
XX
XX 27-JUN-2002.
XX
XX 03-DEC-2001; 2001WO-US45105.
PF
XX 01-DEC-2000; 2000US-250147P.
XX
XX (AMHP) AMERICAN HOME PROD CORP.
XX
XX Young KH, Cao J, Sheu Y D;
PI
XX WPI; 2002-500622/53.
DR
XX Novel cell for detecting ability of test sample to alter RGS
PT protein-mediated reporter gene expression, comprises a heterologous
PT nucleic acid encoding reporter operably linked to pheromone-responsive
PT promoter -
XX
XX Example 7; Page 31; 95pp; English.
XX
XX The invention relates to a cell (I) that responds to a pheromone
CC comprising a heterologous nucleic acid encoding a reporter operably
CC linked to a pheromone-responsive promoter (the reporter is Renilla
CC luciferase, Photinus luciferase, green fluorescent protein (GFP) or its
CC derivative or a heterologous nucleic acid encoding chimeric regulator of
CC G-protein (RGS protein)). (I) is useful for detecting the ability of a
CC test sample to alter RGS protein-mediated reporter gene expression,
CC by: (a) providing (I) where expression of heterologous nucleic acid
CC molecule produces a measurable signal, providing another cell similar to
CC (I) further comprising a second heterologous nucleic acid encoding an RGS
CC protein, incubating a test sample with the cells in the presence of a
CC pheromone under conditions suitable to detect the measurable signal,
CC detecting the level of expression of the heterologous nucleic acid,
CC encoding the reporter and comparing the level of expression of the
CC reporter in the first and second cells, where a difference in the level
CC of expression indicates that the test sample alters RGS protein-mediated
CC reporter gene expression; or (b) providing two aliquots of (I),
CC incubating the aliquots of cells in the presence of a pheromone under
CC conditions suitable to detect the measurable signal, where one of the
CC aliquots contains a test sample, detecting the level of expression of the
CC heterologous nucleic acid encoding the reporter in the aliquots and
CC comparing the level of expression of the reporter in the aliquots, where
CC a difference in the level of expression between the aliquots indicates
CC that the test sample alters RGS protein-mediated reporter gene
CC expression. The test sample used is at a single concentration or in the
CC range of concentrations. The expression of reporter gene is detected in a
CC halo assay, detected spectrometrically or the detection is automated.
CC Since the method of detecting the ability of test sample to modulate RGS
CC mediated reporter gene expression using (I) can be automated, the method
CC can be used for screening test compounds on a large scale. The method is
CC sensitive, can be carried out within a short test period, and is ideal
CC for high through-put systems used in large scale drug discovery. The
CC present sequence is that of a PCR primer used to generate constructs for
CC testing the functional complementation of mammalian RGS proteins in
CC examples of the invention.
XX
XX Sequence 35 BP; 11 A; 13 C; 8 G; 3 T; 0 other;
SQ
Query Match 1.0%; Score 19; DB 24; Length 35;

Best Local Similarity 71.4%; Pred. No. 2.2e+05;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1134 GCTGGCGTTATTGGGATTGTGTGACAAACAGG 1168
|||||
Db 35 GCTGGCGTTGTGTCGGATTGTGTCATCTGGATCCGG 1

RESULT 20
AAZ46427
ID AAZ46427 standard; DNA; 31 BP.
XX AC AAZ46427;
XX DT 06-MAR-2000 (first entry)
XX DE Mutant pACYC184 plasmid constructing linker 184delTet-2.
XX KW Mutational vector; genetic change; therapeutic; mutant; linker; ss.
XX OS Synthetic.
XX PN W0958723-A1.
XX PD 18-NOV-1999.
XX PF 12-MAY-1999; 99WO-US10514.
XX PR 12-MAY-1998; 98US-0078063.
XX PA (KIME-) KIMERAGEN INC.
XX PI Kumar R, Metz RA;
XX DR WPI; 2000-062307/05.
XX PT New mutational vectors, used to introduce specific genetic changes in target DNA sequences in prokaryotic and eukaryotic cells or episomes -
XX PS Examples; Page 13; 37pp; English.
XX CC The invention relates to new mutational vectors that contains a nucleobase strand which alters the sequence of DNA in target cells. The mutational vectors can be used to introduce specific genetic changes in target DNA sequences in prokaryotic and eukaryotic cells or episomes. Such changes can be used to create new phenotypic traits not found in nature, in a subject as a therapeutic or prophylactic intervention and as an investigational tool. Sequences AAZ46426-427 represent linkers replacing a deleted tetracycline region of pACYC184.
XX SQ Sequence 31 BP; 6 A; 7 C; 8 G; 10 T; 0 other;

Query Match 1.0%; Score 18.8; DB 21; Length 31;
Best Local Similarity 76.7%; Pred. No. 2.4e+05;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 852 TAGGTTGCGTTGCTCTCGGTTTGGAGCCT 881
|||||
Db 2 TAGGTTTCACTGCATCTCGAGATTGGATCCT 31

RESULT 21
AAC67280/c
ID AAC67280 standard; DNA; 33 BP.
XX AC AAC67280;
XX DT 09-APR-2001 (first entry)
XX DE Human nuclear receptor NOT1 DNA response element.
XX KW Human; nuclear receptor; NOT1; splice variant; neurological disease; immune disease; cancer; cardiovascular disease; ds.
XX

XX OS Homo sapiens.
XX PN W0200077202-A1.
XX PD 21-DEC-2000.
XX PF 14-JUN-2000; 2000WO-GB02317.
XX PR 15-JUN-1999; 99GB-0013863.
XX PA (SMIK) SMITHKLINE BEECHAM PLC.
XX PI Kremer A, Jackson AR, Cairns WJ;
XX DR WPI; 2001-071271/08.
XX PT Nuclear receptor NOT1 splice variant polypeptide, useful for treating metabolic disorders, cancer, neurological and immune disorders and for identifying agonists/antagonists useful in therapy -
XX PS Example 1; Page 22; 37pp; English.
XX CC The present invention provides the protein and coding sequences of the human nuclear receptor NOT1 splice variant NOT1A. The NOT1A protein is involved in the control of gene expression and the sequences can be used in the treatment of cancer, neurological and immune disorders and cardiovascular disease.
XX SQ Sequence 33 BP; 10 A; 8 C; 9 G; 6 T; 0 other;

Query Match 1.0%; Score 18.8; DB 22; Length 33;
Best Local Similarity 76.7%; Pred. No. 2.4e+05;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1352 ATTCATGACCTTTTGGCTACCTGTGCGAGTC 1381
|||||
Db 32 ATTCATGACCTTTTGGCTGACCTGTGCGATCTC 3

RESULT 22
AAZ80468/c
ID AAZ80468 standard; DNA; 40 BP.
XX AC AAZ80468;
XX DT 23-AUG-1999 (first entry)
XX DE Methanococcus jannaschii DNA polymerase PCR primer #1.
XX KW Methanococcus jannaschii; DNA polymerase; genetic engineering; PCR primer; ss.
XX OS Synthetic.
XX OS Methanococcus jannaschii.
XX PN JP11151087-A.
XX PD 08-JUN-1999.
XX PF 19-NOV-1997; 97JP-0318665.
XX PR 19-NOV-1997; 97JP-0318665.
XX PA (TAKI) TAKARA SHUZO CO LTD.
XX DR WPI; 1999-388477/33.
XX PT New DNA polymerase gene - useful in genetic engineering
XX PS Example 1; Page 13; 19pp; Japanese.
XX CC The present sequence represents a PCR primer for a Methanococcus

CC jannaschii DNA polymerase. The DNA polymerase gene is useful in the
 CC field of gene engineering.

SQ Sequence 40 BP; 19 A; 2 C; 9 G; 10 T; 0 other;
 Query Match 1.0%; Score 18.8; DB 20; Length 40;
 Best Local Similarity 68.4%; Pred. No. 2.7e+05;
 Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1646 TTTAATTTTGCATTATTATGACCATAGGATTTCTTCT 1683
 DB 40 TCTAAGAAATTTATTATTATTCCTCATATGATATCTTCT 3

RESULT 23
 ID ABK59400 standard; RNA; 37 BP.
 XX
 AC ABK59400;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Human CLCA1 gene enzymatic nucleic acid #3771.
 XX
 KW Human; chloride channel calcium activated 1; CLCA1; ss; antiasthmatic;
 KW antiinflammatory; chronic obstructive pulmonary disease; COPD; asthma;
 KW chronic bronchitis; cystic fibrosis; obstructive bowel syndrome;
 KW oxygen therapy; bronchodilator; corticosteroid; vaccination; mucokinetic;
 KW acetylcysteine.
 XX
 OS Homo sapiens.
 XX
 PN WO200211674-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 09-AUG-2001; 2001WO-US24970.
 XX
 PR 09-AUG-2000; 2000US-224383P.
 XX
 PA (RIBO-) RIBOZYME PHARM INC.
 PA (SYNT) SYNTX USA LLC.
 PA (THOM) THOMPSON J.
 XX
 PI Thompson J, McSwiggen J, McKenzie T, Ayers D, Szymkowski DE;
 PI Grupe A;
 XX
 DR WPI; 2002-217145/27.
 XX
 PT Enzymatic polynucleotide that down regulates expression of chloride
 PT channel calcium activated gene, useful for treating Chronic obstructive
 PT pulmonary disease (COPD), chronic bronchitis and asthma -
 XX
 PS Claim 5; Page 95; 152pp; English.
 XX
 CC The invention relates to enzymatic nucleic acid molecules that down
 CC regulate expression of chloride channel calcium activated 1 (CLCA1) genes
 CC by cleaving RNA derived from the genes. The nucleic acid sequences are
 CC useful as pharmaceutical agents for treating conditions such as chronic
 CC obstructive pulmonary disease (COPD), chronic bronchitis, asthma, cystic
 CC fibrosis, obstructive bowel syndrome and any other diseases or conditions
 CC that are related to or will respond to the levels of CLCA1 in a cell or
 CC tissue. The sequences are useful for reducing CLCA1 activity in a cell,
 CC hence, are useful for treatment of a patient having a condition
 CC associated with the level of CLCA1, where the invention further comprises
 CC the use of one or more therapies under conditions suitable for the
 CC treatment, for example, oxygen therapy, bronchodilators, corticosteroids,
 CC antibacterials, vaccinations, acetylcysteine and mucokinetic agents. The
 CC nucleic acids of the invention are also used as diagnostic tools to
 CC examine genetic drift and mutations within diseased cells or to detect
 CC the presence of CLCA1 RNA in a cell. This sequence represents an
 CC enzymatic nucleic acid molecule of the invention.

SQ Sequence 37 BP; 10 A; 6 C; 10 G; 11 U; 0 other;
 Query Match 1.0%; Score 18.6; DB 24; Length 37;
 Best Local Similarity 45.5%; Pred. No. 2.9e+05;
 Matches 15; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

OY 1812 TTTAGTCCCAAGGCCCTGTGTGTTAAATTT 1844
 DB 4 UUCAGGCCGAAGGCGAGGAGGUCUUAACUUU 36

RESULT 24
 ID AAQ31787 standard; DNA; 38 BP.
 XX
 AC AAQ31787;
 XX
 DT 25-MAR-2003 (updated)
 DT 14-APR-1993 (first entry)
 XX
 DE Primer KS4.
 XX
 KW Primer; recombinant DNA; variable domain; light; heavy; chain;
 KW antilysozyme; antibody; D1.3; PCR; polymerase chain reaction;
 KW bridging; linker; flexible; Fv fragment; ss.
 XX
 OS Synthetic.
 XX
 PN WO9220805-A1.
 XX
 PD 26-NOV-1992.
 XX
 PF 11-MAY-1992; 92WO-SE00304.
 XX
 PR 13-MAY-1991; 91SE-0001433.
 XX
 PA (FABR) FABRE MEDICAMENT SA PIERRE.
 XX
 PI Hansson M, Nguyen TN, Nygren PA, Stahl S, Uhlen M;
 PI WPI; 1992-415783/50.
 XX
 DR Recombinant DNA encoding staphylococcal proteins - useful for
 PT isolation and identification of Gram positive bacteria and in
 PT vaccines
 PT
 XX
 PS Disclosure; Page 8; 30pp; English.
 XX
 CC The sequences given in AAQ31784-89 are primers which were used in the
 CC construction of a recombinant DNA sequence containing the two
 CC variable domains of the light and heavy chains of the anti-lysozyme
 CC antibody D1.3. The two variable regions were then linked by PCR via
 CC a 15 amino acid bridging linker which was highly flexible. The
 CC resulting amplification product was a 730 bp gene fragment encoding
 CC a single chain Fv fragment of D1.3, represented by:
 CC NH2-VL-linker-VH-COOH.
 CC (Updated, on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 38 BP; 6 A; 9 C; 9 G; 14 T; 0 other;
 Query Match 1.0%; Score 18.6; DB 13; Length 38;
 Best Local Similarity 72.7%; Pred. No. 2.9e+05;
 Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 848 CTTTATAGTTCGCTCTCGGTTTGAGCC 880
 DB 6 CTTTATAGTTCCTTTTATTCAGCTTGTGTC 38

RESULT 25
 ID AAQ79013 standard; DNA; 38 BP.
 XX

AC AAQ79013;
XX
DT 25-MAR-2003 (updated)
DT 04-AUG-1995 (first entry)
XX
DE SREBP probe #4.
XX
XX Probe; sterol regulatory element binding protein; regulatory protein;
KW cholesterol metabolism; sterol regulatory element-1; SREBP; SRE-1;
KW SREBP-1; SREBP-2; basic-helix-loop-helix-leucine zipper; plasma;
KW transcription factor; low density lipoprotein; LDL; receptor;
KW cholesterol; hypercholesterolaemia; ss.
XX
OS Synthetic.
XX
XX WO9426922-A2.
PN
XX 24-NOV-1994.
PD
XX 13-MAY-1994; 94WO-US05300.
PF
XX 13-MAY-1993; 93US-0061697.
PR
PR 01-OCT-1993; 93US-0131365.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
PA
XX
XX Briggs MR, Brown MS, Goldstein JL, Wang X;
PI WPI; 1995-006813/01.
DR
XX
XX New sterol regulator element binding protein - used to develop
PT prods. and screening assays for agents for reducing plasma
PT cholesterol levels (Eng)
XX
XX Disclosure; Page 182; 305pp; English.
PS
XX
XX The sequences given in AAQ79010-15 are probes which were used in the
CC isolation of recombinant DNA encoding the sterol regulatory element
CC binding protein (SREBP) of the invention. SREBP's are regulatory
CC proteins which are involved in the regulation of genes involved in
CC cholesterol metabolism that are under the control of an associated
CC sterol regulatory element-1 (SRE-1). SREBP proteins fall into two
CC families, SREBP-1 and SREBP-2. Each have the ability to bind to SRE
CC sequences and modulate SRE-mediated transcription. Both proteins
CC are members of a family of basic-helix-loop-helix-leucine zipper
CC transcription factors. SREBP promotes SRE-1-mediated gene
CC transcription, eg. low density lipoprotein (LDL) receptor production
CC in the presence of sterols. SREBP identified in screening assays, may
CC be used to reduce plasma cholesterol levels and in controlling
CC hypercholesterolaemia and its associated diseases.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 38 BP; 7 A; 4 C; 14 G; 13 T; 0 other;
Query Match 1.0%; Score 18.6; DB 16; Length 38;
Best Local Similarity 72.7%; Pred. No. 2.9e+05;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 121 AAACGTCCTTGAAGAAATCATCCAGTGTC 153
Db
35 AATCACCCTCTGTAATAATCACCCTGTC 3
RESULT 26
ABX67373/c
ID ABX67373 standard; DNA; 30 BP.
XX
AC ABX67373;
XX
XX 07-MAY-2003 (first entry)
DT
XX Novel Helicobacter pylori gene PCR primer #344.
XX

KW Protein-protein interaction; ulcer; selected interacting domain;
KW SID; PCR; primer; ss.
XX
OS Helicobacter pylori.
XX
PN WO200266501-A2.
XX
XX 29-AUG-2002.
PD
XX 28-DEC-2001; 2001WO-EP15428.
PF
XX 02-JAN-2001; 2001US-259302P.
PR
XX (HYBR-) HYBRIGENICS.
PA (INSP) INST PASTEUR.
XX
XX Legrain P, Rain J, Colland F, De Reuse H, Labigne A;
PI WPI; 2002-674910/72.
DR
XX New complexes of protein-protein interactions in Helicobacter pylori,
PT useful for identifying modulating compounds for treating or preventing
PT ulcers in mammals -
XX
XX Example 9; Page 498; 642pp; English.
XX
XX The invention describes a complex of protein-protein interactions in
CC Helicobacter pylori selected from 421 complexes given in the
CC specification. The complex of protein-protein interactions are useful
CC for screening for agents which modulate the interaction of proteins.
CC Modulating compounds which binds to a targeted bacterial protein may be
CC used for treating or preventing ulcers in a human or animal. This
CC sequence represents a primer used to isolate polynucleotides encoding
CC Helicobacter pylori proteins for studies on protein-protein
CC interactions.
XX
SQ Sequence 30 BP; 15 A; 3 C; 6 G; 3 T; 3 U; 0 other;
Query Match 1.0%; Score 18.4; DB 24; Length 30;
Best Local Similarity 78.6%; Pred. No. 3e+05; 6; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 6;
QY 1640 TCTGCTTTTAATTGCAATTATTATGA 1667
Db 30 TCTTCAATTAATTTCCCTCATGATGA 3
RESULT 27
ABL41520/c
ID ABL41520 standard; DNA; 33 BP.
XX
AC ABL41520;
XX
XX 01-JUL-2002 (first entry)
DT
XX Primer #4 related to human semaphorin protein 9.
DE
XX Human; semaphorin protein 9; nontropic; neuroprotective;
KW nervous system; ss; PCR primer.
XX
XX Homo sapiens.
OS
XX WO200204505-A1.
PN
XX 17-JAN-2002.
PD
XX 18-JUN-2001; 2001WO-CN00995.
PF
XX 19-JUN-2000; 2000CN-0116602.
PR
XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
PA
XX Mao Y, Xie Y;
PI

DT 27-OCT-1998 (first entry)
 XX 5' PCR primer used in the course of the invention.
 DE protein kinase activity; Rho protein; preparation;
 XX therapeutic composition; PCR primer; ss.
 KW Synthetic.
 XX JP10191985-A.
 XX 28-JUL-1998.
 XX 17-JAN-1997; 97JP-0019870.
 XX 17-JAN-1997; 97JP-0019870.
 XX (KIRI) KIRIN BREWERY KK.
 XX WPI; 1998-460110/40.
 XX New protein exhibiting protein kinase activity - is not capable of
 PT binding to active Rho protein or its derivative, used, e.g.
 PT therapeutically
 XX Example 9; Page 20; 60pp; Japanese.
 PS PCR primers AAV43409-30 are used in the course of the invention. The
 CC specification describes a protein which exhibits protein kinase
 CC activity and is not capable of binding to active Rho protein or its
 CC derivative. The materials may be used for the preparation of
 CC therapeutic compositions.
 XX Sequence 38 BP; 16 A; 6 C; 7 G; 9 T; 0 other;
 SQ

Query Match 1.0%; Score 18.4; DB 19; Length 38;
 Best Local Similarity 69.4%; Pred. No. 3.3e+05;
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 570 CTTTGGAGCCATTTTGTGATCTCACTGACTACAG 605
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 36 CTTTGGATTCATTATGCTGCAAGCTTGTAAACAG 1

RESULT 31
 AAX80470
 ID AAX80470 standard; DNA; 40 BP.
 AC AAX80470;
 XX 23-AUG-1999 (first entry)
 DT Methanococcus jannaschii DNA polymerase PCR primer #3.
 XX Methanococcus jannaschii; DNA polymerase; genetic engineering;
 KW PCR primer; ss.
 XX Synthetic.
 OS Methanococcus jannaschii.
 XX JP11151087-A.
 XX 08-JUN-1999.
 XX 19-NOV-1997; 97JP-0318665.
 XX 19-NOV-1997; 97JP-0318665.
 XX (TAKI) TAKARA SHUZO CO LTD.
 XX WPI; 1999-388477/33.
 XX New DNA polymerase gene - useful in genetic engineering

XX Example 1; Page 13; 19pp; Japanese.
 PS The present sequence represents a PCR primer for a Methanococcus
 CC jannaschii DNA polymerase. The DNA polymerase gene is useful in the
 CC field of gene engineering.
 XX Sequence 40 BP; 9 A; 5 C; 10 G; 16 T; 0 other;
 SQ

Query Match 1.0%; Score 18.4; DB 20; Length 40;
 Best Local Similarity 69.4%; Pred. No. 3.4e+05;
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1519 CCTGTTTACCCTCATGCTGTTTGTAGTTAATGATGC 1554
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2 CCGATTTAAGCATATGATTGTTATGTTCAATGTTGC 37

RESULT 32
 AAC60772
 ID AAC60772 standard; DNA; 40 BP.
 XX AAC60772;
 AC AAC60772;
 XX 06-FEB-2001 (first entry)
 DT Vitamin D 24 hydroxylase (CYP24) forward PCR primer SEQ ID NO:1.
 XX Vitamin D 24 hydroxylase; vitamin D receptor; CYP24; VDR; ZNF24;
 KW oncogene; breast cancer; chromosome 20; 20q13.2; PCR primer;
 KW 25-hydroxyvitamin D3 24-hydroxylase enzyme; cytosolic; ss.
 XX Homo sapiens.
 OS WO200060109-A1.
 XX 12-OCT-2000.
 XX 06-MAR-2000; 2000WO-US05972.
 XX 02-APR-1999; 99US-0285292.
 XX (REGC) UNIV CALIFORNIA.
 XX Albertson DG, Pinkel D, Collins C, Gray JW, Ystra B;
 XX WPI; 2000-656233/63.
 XX Detecting a predisposition to or a progression of cancer especially
 PT breast cancer in humans comprises detecting levels of CYP24 in a
 PT biological sample -
 XX Example 1; Page 53; 73pp; English.
 PS The present invention describes a method for detecting (I) a
 CC predisposition to cancer in an animal. The method comprises detecting
 CC the level of CYP24 (25-hydroxyvitamin D3 24-hydroxylase enzyme) in a
 CC biological sample from the animal and comparing it with a control sample
 CC taken from a normal, cancer-free tissue, where an increased level of
 CC CYP24 in the biological sample compared to the control sample indicates
 CC a predisposition to cancer in the animal. (I) is useful for detecting a
 CC predisposition to cancer in humans, non-human primates, canines, felines,
 CC murines, bovines, equines, porcines and lagomorphs. An example from the
 CC present invention describes the identification of CYP24 as a driver
 CC oncogene, for amplification at chromosome band 20q13.2; the present
 CC sequence represents a PCR primer for CYP24 which is used in this example.
 XX Sequence 40 BP; 14 A; 11 C; 8 G; 7 T; 0 other;
 SQ

Query Match 1.0%; Score 18.4; DB 21; Length 40;
 Best Local Similarity 69.4%; Pred. No. 3.4e+05;
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

PR 15-JUN-2001; 2001US-0212308.
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX Lyamichev V, Allawi H, Dong F, Neri BP, Vener IT;
PI WPI; 2002-049698/06.
XX
XX Identifying oligonucleotides hybridizing to nucleic acids containing
PT secondary structure, useful in clinical diagnosis, comprises
PT identifying primers that interact with the target to form an extension
PT product under amplification conditions.
XX
XX Example 18; Fig 54A; 409pp; English.
XX
XX The present invention describes a method for identifying oligonucleotides
CC with desired hybridisation properties to nucleic acid targets containing
CC secondary structure. The method comprises amplifying a target nucleic
CC acid having at least one accessible and one inaccessible site. Primers
CC that form an extension product are identified as the oligonucleotides
CC which can interact with the folded target nucleic acid. Oligonucleotides
CC from the present invention can be used in novel detection methods for
CC clinical diagnostic purposes, including the detection and identification
CC of pathogenic organisms (e.g. HIV). The method allows the ability to
CC rapidly analyse nucleic acid structures. ABL46034 to ABL46367 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 33 BP; 17 A; 0 C; 7 G; 9 T; 0 other;
SQ

Query Match 1.0%; Score 18.2; DB 24; Length 33;
Best Local Similarity 74.2%; Pred. No. 3.5e+05;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 531 AAGTCTTAAAAAATCTGTAGAAAAAGGATA 561
DB 2 AAGTTTTTAAAAAGTTGAAGTAAAAAGGAGA 32

RESULT 36
AAV37858/c
ID AAV37858 standard; DNA; 36 BP.
XX
XX AAV37858;
XX
XX 10-SEP-1998 (first entry)
XX
XX PEA toxin gene and sfv-Lym PCR primer SEQ ID NO:28.
DE
XX Immunotoxin; tumour; HIV antigen; lymphocyte; neuron; cytotoxic cell;
KW cancer; PCR primer; ss.
XX
XX Synthetic.
XX Homo sapiens.
XX WO9817116-A1.
XX
XX 30-APR-1998.
XX
XX 23-OCT-1997; 97WO-US19206.
XX
XX 23-OCT-1996; 96US-0740003.
XX
XX (UYWA-) UNIV WAKE FOREST.
XX
XX Chen S;
XX
XX WPI; 1998-271753/24.
XX
XX Cytotoxic lymphocytes or neurons expressing immunotoxin - useful
KW for the treatment of cancer or virally-infected cells
XX
XX Synthetic.
XX Homo sapiens.
XX WO9817116-A1.
XX
XX 30-APR-1998.
XX
XX 23-OCT-1997; 97WO-US19206.
XX
XX 23-OCT-1996; 96US-0740003.
XX
XX (UYWA-) UNIV WAKE FOREST.
XX
XX Chen S;
XX
XX WPI; 1998-271753/24.
XX
XX Cytotoxic lymphocytes or neurons expressing immunotoxin
PT for the treatment of cancer or virally-infected cells
XX
XX Example 4; Page 45; 82pp; English.
XX

CC The present sequence represents a PCR primer, from the present
CC invention, used for amplifying the PEA toxin gene and sfv-Lym.
CC The present invention describes cytotoxic lymphocytes or neurons
CC that express and secrete an immunotoxin. The immunotoxin secreted
CC by the cell is used to kill cancer, especially breast, ovarian,
CC gastric or brain cancer and virally-infected cells, especially
CC in HIV infected patients.
XX
XX Sequence 36 BP; 8 A; 10 C; 10 G; 8 T; 0 other;
SQ

Query Match 1.0%; Score 18.2; DB 19; Length 36;
Best Local Similarity 74.2%; Pred. No. 3.7e+05;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 998 GAAGCTGCTCTTCTTCTAGAGACTTCGGTA 1028
DB 33 GAAGCTGAGTCGCTCTAGAGAAATTCGGTA 3

RESULT 37
AAV37850/c
ID AAV37850 standard; DNA; 36 BP.
XX
XX AAV37850;
XX
XX 10-SEP-1998 (first entry)
XX
XX Transformed lymphocyte, MOLT-2Fv23e-PE40, PCR primer SEQ ID NO:20.
DE
XX Immunotoxin; tumour; HIV antigen; lymphocyte; neuron; cytotoxic cell;
KW cancer; PCR primer; ss.
XX
XX Synthetic.
XX Homo sapiens.
XX WO9817116-A1.
XX
XX 30-APR-1998.
XX
XX 23-OCT-1997; 97WO-US19206.
XX
XX 23-OCT-1996; 96US-0740003.
XX
XX (UYWA-) UNIV WAKE FOREST.
XX
XX Chen S;
XX
XX WPI; 1998-271753/24.
XX
XX Cytotoxic lymphocytes or neurons expressing immunotoxin - useful
PT for the treatment of cancer or virally-infected cells
XX
XX Example 1; Page 40; 82pp; English.
XX
XX The present sequence represents a PCR primer, from the present
CC invention, used for amplifying the transformed lymphocyte.
CC MOLT-2Fv23e-PE40. The present invention describes cytotoxic lymphocytes
CC or neurons that express and secrete an immunotoxin. The immunotoxin
CC secreted by the cell is used to kill cancer, especially breast,
CC ovarian, gastric or brain cancer and virally-infected cells,
CC especially in HIV infected patients.
XX
XX Sequence 36 BP; 8 A; 10 C; 10 G; 8 T; 0 other;
SQ

Query Match 1.0%; Score 18.2; DB 19; Length 36;
Best Local Similarity 74.2%; Pred. No. 3.7e+05;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 998 GAAGCTGCTCTTCTTCTAGAGACTTCGGTA 1028
DB 33 GAAGCTGAGTCGCTCTAGAGAAATTCGGTA 3

RESULT 38
 ABZ58980
 ID ABZ58980 standard; DNA; 39 BP.
 XX
 AC ABZ58980;
 XX
 DT 28-APR-2003 (first entry)
 XX
 DE End-locked five-helix protein gene constructing primer P4.
 XX
 KW Five-helix protein; HIV; gp41; virucide; anti-HIV; vaccine; PCR;
 KW immune response; viral membrane; Ebola viral infection; primer; ss.
 XX
 OS Synthetic.
 XX
 PN WO2003006056-A2.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2002; 2002WO-US21915.
 XX
 PR 11-JUL-2001; 2001US-304152P.
 PR 11-JUL-2002; 2002US-0193412.
 XX
 PA (ZHOU/) ZHOU G.
 XX
 PI Zhou G;
 XX
 DR WPI; 2003-221659/21.
 XX
 PT New end-locked five-helix protein comprising 3 N-helices and 2
 PT C-helices of HIV gp41, 4 inside linkers, and at least one terminal
 PT linker, useful for inhibiting HIV infection in a human, and for
 PT eliciting an immune response to HIV in a host -
 XX
 PS Example 1; Page 18; 21pp; English.
 XX
 CC The invention relates to new end-locked five-helix protein comprising 3
 CC N-helices and 2 C-helices of HIV gp41, 4 inside linkers, and at least
 CC one terminal linker. The helices are connected by the inside linker, and
 CC the terminal linker is connected to a helix and is capable of cross-
 CC linking with one of the inside linkers. The end-locked five-helix protein
 CC is useful for inhibiting entry of HIV into a cell, for inhibiting HIV
 CC infection in a human, and for eliciting an immune response to HIV in a
 CC host. The proteins may also be used as fusion inhibitors to block viral
 CC entry or viral membrane fusion, as vaccine (e.g. against Ebola viral
 CC infection), as tools for screening drug candidates against HIV, and as
 CC inhibitors for blocking Ebola viruses. Sequences ABZ58977-988 represent
 CC PCR primers used for constructing the gene encoding the end-locked five-
 CC helix proteins of the invention.
 XX
 SQ Sequence 39 BP; 10 A; 14 C; 4 G; 11 T; 0 other;
 Query Match 1.0%; Score 18.2; DB 25; Length 39;
 Best Local Similarity 74.2%; Pred. No. 3.8e+05;
 Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1587 CAACACCTCAACTCTATCAGTTTCTTAAAT 1617
 Db | | | | | | | | | | | | | | | | | | | |
 8 CCACAGCCACCACTCAATAATCTTGTTCAT 38
 RESULT 39
 AAX27242
 ID AAX27242 standard; DNA; 40 BP.
 XX
 AC AAX27242;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE PCR primer for Tetanus toxin fragment C coding sequence.
 XX
 KW Tetanus toxin fragment C; TTC; central nervous system; CNS; spinal cord;

KW proteolytic fragment; retrograde axonal transport; spinal cord disease;
 KW transynaptic transport; neurodegenerative disease; motoneuron disease;
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; therapy; ALS;
 KW SMA; neurodegenerative lysosomal storage disease; neuronal mapping;
 KW PCR primer; ss.
 XX
 OS Synthetic.
 OS Clostridium tetani.
 XX
 PN WO9909057-A2.
 XX
 PD 25-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-EP05113.
 XX
 PR 13-NOV-1997; 97US-0065236.
 PR 14-AUG-1997; 97US-0055615.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Brulet P, Coen L, Osta Pinzolas R;
 XX
 DR WPI; 1999-180971/15.
 XX
 PT Delivery of a composition to the central nervous system or spinal
 PT cord - comprises administration of a non-toxic, proteolytic
 PT fragment of tetanus toxin in association with a molecule having
 PT biological function
 XX
 PS Example 1; Page 19; 53pp; English.
 XX
 CC This sequence represents a PCR primer for DNA encoding the tetanus toxin
 CC fragment C (TTC). The invention relates to a method for in vivo delivery
 CC of a desired composition into a human or animal central nervous system
 CC (CNS) or spinal cord comprising administering a non-toxic, proteolytic
 CC fragment of tetanus toxin (TT) in association with at least a molecule
 CC having a biological function and where the composition is capable of in
 CC vivo retrograde axonal transport and transynaptic transport into the CNS
 CC or the spinal cord of the human or animal and of being delivered to
 CC different areas of the CNS or the spinal cord. The method can be used for
 CC the treatment of humans or animals with CNS or spinal cord disease,
 CC e.g. neurodegenerative and motoneuron diseases such as amyotrophic
 CC lateral sclerosis (ALS), spinal muscular atrophies (SMA) or
 CC neurodegenerative lysosomal storage diseases. Compositions comprising
 CC hybrid fragments of TT comprising fragments C and B can also be used for
 CC neuronal mapping and immunisations. Use of TT comprising fragments A, B
 CC and C results in better transport of the fragment inside the organism
 CC compared with fragment C.
 XX
 SQ Sequence 40 BP; 9 A; 11 C; 10 G; 10 T; 0 other;
 Query Match 1.0%; Score 18.2; DB 20; Length 40;
 Best Local Similarity 74.2%; Pred. No. 3.9e+05;
 Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1376 GCAGTCTGTTTATAGACACATCTTTTGA 1406
 Db | | | | | | | | | | | | | | | | | | | |
 2 GCAGTCTGAGTCTAGACCATGGCTTTTGA 32
 RESULT 40
 AAL46329/c
 ID AAL46329 standard; DNA; 28 BP.
 XX
 AC AAL46329;
 XX
 DT 19-JUL-2002 (first entry)
 XX
 DE M30 protein vector PCR primer SEQ ID NO: 23.
 XX
 KW Neurodegenerative disease; M30; M31; M32; M33; stroke;
 KW fragile X syndrome; Huntington's disease; Parkinson's disease;
 KW Alzheimer's disease; multiple sclerosis; ovarian cancer;

KW neurodegeneration; immune disorder; autoimmune disease; allergy;
 KW infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
 KW immunosuppressive; cytostatic; nootropic; antiparkinsonian; anti-allergic;
 KW virucide; anti-inflammatory; PCR; primer; ss.
 XX Unidentified.
 OS
 XX WO20021138-A2.
 PN
 XX
 XX
 PD 14-MAR-2002.
 XX
 XX
 PF 07-SEP-2001; 2001WO-EP10366.
 XX
 XX 07-SEP-2000; 2000US-0657479.
 PR
 XX (AXAR-) AXARON BIOSCIENCE AG.
 PA
 XX
 XX Schneider A, Hiemisch H, Rossner M, Klugmann M, Naim J;
 PI Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D;
 PI Scheek S;
 XX
 XX WPI; 2002-292287/33.
 DR
 XX
 XX Diagnosis of neurodegenerative disease comprises detecting level of
 PT M30-family proteins -
 PT
 XX
 XX Example 14; Page 121; 130pp; German.
 PS
 XX
 XX The present invention relates to a method of diagnosing neurodegenerative
 CC diseases, comprising determining the concentration of a protein in a body
 CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
 CC M33. The method is used to diagnose neurodegenerative diseases,
 CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
 CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
 CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
 CC especially ovarian cancer. The proteins can be used to identify specific
 CC ligands, potentially useful for treating neurodegeneration, immune-system
 CC disorders (e.g. autoimmune diseases, allergy, viral infection,
 CC leukaemia, inflammation etc.), carcinoma and sarcoma. Inhibitors of the
 CC interaction between the proteins and the protein kinase IRAK-1 can be
 CC used to treat neurodegeneration. The present sequence is a PCR primer for
 CC a coding sequence of a protein used in the method of the invention.
 XX
 SQ Sequence 28 BP; 8 A; 7 C; 7 G; 6 T; 0 other;
 Query Match 1.0%; Score 18; DB 24; Length 28;
 Best Local Similarity 80.8%; Pred. No. 3.7e+05;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 723 TCTTCATGCTGTCGATGAATGACC 748
 Db 26 TCTTCATGCTGGGAGAGATTCGATC 1
 RESULT 41
 AAI68954/c
 ID AAI68954 standard; DNA; 32 BP.
 XX
 AC AAI68954;
 XX
 XX 29-JAN-2002 (first entry)
 DT
 XX
 DE Activated T-cell derived DNA fragment, #90.
 XX
 KW Activated T-cell; immunosuppressive; immunostimulant; anti-inflammatory;
 KW cytostatic; gene therapy; vaccine; allergen; transplant rejection;
 KW guest versus host disease; malignant disease; ds.
 XX
 XX Homo sapiens.
 OS
 XX DE10021834-A1.
 PN
 XX 15-NOV-2001.
 PD

XX
 PF 06-MAY-2000; 2000DE-1021834.
 XX
 PR 06-MAY-2000; 2000DE-1021834.
 XX
 PA (LYNX-) LYNX THERAPEUTICS GMBH.
 XX
 XX Kramer MD, Winter H, Reinhartz J;
 XX
 XX WPI; 2002-027320/04.
 DR
 XX
 XX New mRNA indicative of T cell activation and functional status, useful
 PT for diagnosis and therapy e.g. of autoimmunity or transplant rejection
 PT
 XX
 XX Claim 1; Page 30; 94pp; German.
 PS
 XX
 XX This sequence represents a novel messenger RNA, (mRNA), (I), for use as
 CC indicator of the activation and functional status of T cells, that have
 CC increased or reduced expression, and are present at higher or lower
 CC concentration, in activated T cells, relative to normal or resting cells,
 CC where (I) hybridizes to any of 334 sequences, reproduced, or their
 CC derivatives, complements or fragments. The products of the invention
 CC have immunosuppressive, immunostimulant, anti-inflammatory and cytostatic
 CC activity and can be used for gene therapy. The polynucleotides of the
 CC invention are used: (i) as reagent for detecting activation/functional
 CC status of T cells, for diagnosis, therapy, modulation or control of the
 CC status, in cases of (auto)immunity (against microorganisms, vaccines or
 CC allergens); transplant rejection; immunologically-related inflammation;
 CC immunosuppression; immune deficiency; guest versus host disease, and
 CC malignant diseases of the immune system; (ii) for identifying agents,
 CC potential pharmaceuticals, that bind to (II) or derived polypeptides
 CC (III); (iii) to prepare kits for measuring gene expression profiles in
 CC isolated immune, especially T, cells; (iv) to raise antibodies (Ab)
 CC directed against (III); and (v) to prepare binding molecules (IV)
 CC specific for (III). Ab and (iv) are also useful for detecting and
 CC modulating the activation and functional status of T cells.
 CC AAI68865-AAI69198 represent the activated T-cell derived polynucleotide
 CC fragments described in the method of the invention.
 XX
 SQ Sequence 32 BP; 14 A; 6 C; 6 G; 6 T; 0 other;
 Query Match 1.0%; Score 18; DB 24; Length 32;
 Best Local Similarity 80.8%; Pred. No. 4e+05;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1592 GCTACACTCTATCATGTTTGTGTAAT 1617
 Db 27 GCTAGCAGTCTATCAATTTTGTGTAT 2
 RESULT 42
 ABV72966
 ID ABV72966 standard; DNA; 33 BP.
 XX
 AC ABV72966;
 XX
 XX 08-JAN-2003 (first entry)
 DT
 XX
 DE Human PWMP protein 60 cDNA cloning PCR primer 3.
 XX
 KW PWMP protein 60; tumour; nosohemia; HIV infection; immunological disease;
 KW inflammation; human; PCR; primer; ss.
 XX
 XX Homo sapiens.
 OS
 XX CN1307055-A.
 PN
 XX
 PD 08-AUG-2001.
 XX
 XX 28-JAN-2000; 2000CN-0111613.
 PF
 XX 28-JAN-2000; 2000CN-0111613.
 PD

XX PA (BODA-) BODAO GENE TECH CO LTD SHANGHAI.
 XX PI Mao Y, Xie Y;
 XX WPI; 2002-049884/07.
 XX
 XX New polypeptide, useful for treating diseases, comprises the human PWMP
 PT protein 60 and polynucleotide for coding said polypeptide -
 XX
 XX Example 5; Page 17 (disclosure); 34pp; Chinese.
 XX
 CC The invention relates to a novel human PWMP protein 60 polypeptide, and
 CC encoding polynucleotides. The polypeptide can be expressed by standard
 CC DNA recombination. The PWMP protein 60 polypeptide, polynucleotide and
 CC modulators are useful for treating various diseases, such as malignant
 CC tumour, nosohemia, HIV infection, immunological diseases and
 CC inflammations. The present sequence represents the human PWMP protein 60
 CC cDNA cloning PCR primer.
 XX
 SQ Sequence 33 BP; 12 A; 8 C; 3 G; 10 T; 0 other;
 Query Match 1.0%; Score 18; DB 24; Length 33;
 Best Local Similarity 80.8%; Pred. No. 4e+05;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1455 TGGGGTATTTCATTAAACATACTTTG 1480
 DB 8 TGATGTCATCAATAACATACTTTG 33
 RESULT 43
 AAS44188/c
 ID AAS44188 standard; DNA; 34 BP.
 XX
 AC AAS44188;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Neisseria meningitidis B MC58 genomic DNA sequence PCR primer #283.
 XX
 KW Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ss;
 KW Neisserial protein; PCR primer.
 XX
 OS Neisseria meningitidis.
 OS Synthetic.
 XX
 PN WO200164922-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-IB00452.
 XX
 PR 28-FEB-2000; 2000GB-0004695.
 PR 13-NOV-2000; 2000GB-0027675.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Arico MB, Comanducci M, Galeotti C, Massignani V, Guiliani MM;
 PI Pizza M;
 XX
 WPI; 2001-582163/65.
 XX
 PT Producing heterologous proteins from Neisseria meningitidis and N.
 PT gonorrhoeae -
 XX
 PS Disclosure; Page 87; 119pp; English.
 XX
 CC The invention relates to methods for the heterologous expression of
 CC Neisserial proteins from Neisseria meningitidis and Neisseria
 CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the
 CC leader peptide, and may be replaced by a domain from a different protein
 CC to make a fusion protein, in order to enhance heterologous expression of

CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine
 CC stretch, can be mutated to enhance expression. The proteins used in the
 CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
 CC AAS43807-AAS43867 and AAS43906-AAS44358 represent PCR primers used in the
 CC methods of the invention.
 XX
 SQ Sequence 34 BP; 13 A; 9 C; 8 G; 4 T; 0 other;
 Query Match 1.0%; Score 18; DB 22; Length 34;
 Best Local Similarity 70.6%; Pred. No. 4.1e+05;
 Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 443 CTCCTCGGATGTTCTTCTATGGGATGCGG 476
 DB 34 CTTTCTCGGATGTTCTTCTATGGGATGCGG 1
 RESULT 44
 ABX05700
 ID ABX05700 standard; DNA; 36 BP.
 XX
 AC ABX05700;
 XX
 DT 11-FEB-2003 (first entry)
 XX
 DE PCR primer for knock out of S. pneumoniae gene SEQ ID 3480 #4.
 XX
 KW PCR; ss; bacterial meningitis; pneumonia; sepsis; otitis media;
 KW ear infection; antiinflammatory; antibacterial; immunostimulant;
 KW auditory; respiratory; gene therapy; vaccine; primer.
 XX
 OS Streptococcus pneumoniae type 4 strain.
 XX
 PN WO200277021-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-IB02163.
 XX
 PR 27-MAR-2001; 2001GB-0007658.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Massignani V, Tettelin H, Fraser C;
 XX
 WPI; 2003-040579/03.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
 PT or ear infection -
 XX
 PS Disclosure; Page 42; 56pp; English.
 XX
 CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS56454. Also included are an antibody which binds one of the
 CC proteins, treating a patient by administering the protein, DNA or
 CC antibody (in a composition), a kit comprising first and second primers,
 CC which are the nucleic acid cited above or fragments between nucleotides
 CC 8-100 of a sequence not defined in the specification, for amplifying a
 CC target sequence contained within a Streptococcus nucleic acid sequence,
 CC where the first primer is substantially complementary to the target
 CC sequence and the second primer is substantially complementary to the
 CC complement of the target sequence, and where the parts of the primers
 CC having substantial complementarity define the termini of the target
 CC sequence to be amplified, assay comprising contacting a test compound
 CC with the protein, and determining whether the test compound binds to the
 CC protein and a Streptococcus pneumoniae bacterium, where one or more

CC genes encoding the proteins has been rendered inactive. The proteins,
 CC nucleic acid molecules, antibody and compositions are useful as
 CC medicaments for treating or preventing a disease or infection due to
 CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
 CC sepsis, otitis media or ear infection. They are also useful in developing
 CC vaccines, diagnostics and antibiotics. The methods are useful for
 CC identifying immunodominant proteins. The present sequence is a PCR
 CC primer used in a knock-out experiment for one of the 2489 identified
 CC coding region from the genomic sequence.

XX Sequence 36 BP; 3 A; 8 C; 7 G; 18 T; 0 other;

Query Match 1.0%; Score 18; DB 25; Length 36;
 Best Local Similarity 70.6%; Pred. No. 4.2e+05;
 Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 360 TCCTTCGGGAGCTGTTCTTCTGAAATCTTTTGCT 393
 |||||
 Db 3 TAGCGGCGCAATGTATTTCTGCTTCTTTTGCT 36

RESULT 45
 ABK58002/c
 ID ABK58002 standard; RNA; 38 BP.
 XX AC ABK58002;
 XX DT 02-JUL-2002 (first entry)
 XX DE Human CLCA1 gene enzymatic nucleic acid #2373.
 XX KW Human; chloride channel calcium activated 1; CLCA1; ss; antiasthmatic;
 KW antiinflammatory; chronic obstructive pulmonary disease; COPD; asthma;
 KW chronic bronchitis; cystic fibrosis; obstructive bowel syndrome;
 KW oxygen therapy; bronchodilator; corticosteroid; vaccination; mucokinetic;
 KW acetylcysteine.
 XX OS Homo sapiens.
 XX PN WO200211674-A2.
 XX PD 14-FEB-2002.
 XX PF 09-AUG-2001; 2001WO-US24970.
 XX PR 09-AUG-2000; 2000US-224383P.
 XX PA (RIBO-) RIBOZYME PHARM INC.
 PA (SYNT) SYNTEX USA LLC.
 PA (THOM/) THOMPSON J.

PI Thompson J, McSwiggen J, McKenzie T, Ayers D, Szymkowski DE;
 PI Grupe A;
 XX WPI; 2002-217145/27.
 XX Enzymatic polynucleotide that down regulates expression of chloride
 PT channel calcium activated gene, useful for treating Chronic obstructive
 PT pulmonary disease (COPD), chronic bronchitis and asthma -
 XX Claim 5; Page 56; 152pp; English.

XX The invention relates to enzymatic nucleic acid molecules that down
 CC regulate expression of chloride channel calcium activated 1 (CLCA1) genes
 CC by cleaving RNA derived from the genes. The nucleic acid sequences are
 CC useful as pharmaceutical agents for treating conditions such as chronic
 CC obstructive pulmonary disease (COPD), chronic bronchitis, asthma, cystic
 CC fibrosis, obstructive bowel syndrome and any other diseases or conditions
 CC that are related to or will respond to the levels of CLCA1 in a cell or
 CC tissue. The sequences are useful for reducing CLCA1 activity in a cell,
 CC hence, are useful for treatment of a patient having a condition
 CC associated with the level of CLCA1, where the invention further comprises
 CC the use of one or more therapies under conditions suitable for the

CC treatment, for example, oxygen therapy, bronchodilators, corticosteroids,
 CC antibacterials, vaccinations, acetylcysteine and mucokinetic agents. The
 CC nucleic acids of the invention are also used as diagnostic tools to
 CC examine genetic drift and mutations within diseased cells or to detect
 CC the presence of CLCA1 RNA in a cell. This sequence represents an
 CC enzymatic nucleic acid molecule of the invention.

XX Sequence 38 BP; 8 A; 8 C; 10 G; 12 U; 0 other;

Query Match 1.0%; Score 18; DB 24; Length 38;
 Best Local Similarity 70.6%; Pred. No. 4.3e+05;
 Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1808 GGACTTTAGTCCAAAGGCCCTGTGTATTATAA 1841
 |||||
 Db 35 GGACTTTCGGCCTAACGGCCTCATCATGATGAAA 2

Search completed: October 23, 2003, 09:41:44
 Job time : 500 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 09:33:14 ; Search time 508 Seconds
(without alignments)
9840.309 Million cell updates/sec

Title: US-09-868-987-1

Perfect score: 1864

Sequence: 1 atggacttcgcgcatattgtc.....atgcacaaatcagataatgc 1864

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 1060730

Minimum DB seq length: 10

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.4	1.1	40	10	US-09-780-929-53
2	19	1.0	35	9	US-09-152-059-56
3	19	1.0	35	12	US-10-008-029-56
4	19	1.0	35	12	US-10-208-650-56
5	18.8	1.0	31	10	US-09-825-124B-12
6	18.8	1.0	39	10	US-09-974-973-7
7	18.6	1.0	25	14	US-10-098-263B-33220
8	18.2	1.0	33	12	US-09-882-945A-151
9	18.2	1.0	39	14	US-10-193-412A-9
10	18.2	1.0	40	11	US-09-816-467-13
11	17.8	1.0	31	9	US-09-801-274-1689
12	17.8	1.0	35	12	US-10-292-896-25
13	17.8	1.0	38	11	US-09-877-478-2673
14	17.8	1.0	38	11	US-09-776-474-2153
15	17.8	1.0	39	9	US-09-732-914-119
16	17.8	1.0	39	10	US-09-922-261-132

17	17.8	1.0	40	11	US-09-738-444A-1	Sequence 1, Appli
18	17.6	0.9	33	10	US-09-805-761-28	Sequence 28, Appl
19	17.4	0.9	25	12	US-10-083-246A-162	Sequence 162, App
20	17.4	0.9	35	9	US-09-861-696-46	Sequence 46, Appl
21	17.4	0.9	35	10	US-09-454-099A-46	Sequence 46, Appl
22	17.4	0.9	40	12	US-10-157-305A-89	Sequence 89, Appl
23	17.4	0.9	40	12	US-10-157-391-89	Sequence 89, Appl
24	17.4	0.9	40	12	US-10-157-096-89	Sequence 89, Appl
25	17.4	0.9	40	12	US-10-157-215A-89	Sequence 89, Appl
26	17.4	0.9	40	12	US-10-157-302-89	Sequence 89, Appl
27	17.4	0.9	40	12	US-10-154-951B-89	Sequence 89, Appl
28	17.4	0.9	40	12	US-10-157-299-89	Sequence 89, Appl
29	17.4	0.9	40	14	US-10-184-007A-2	Sequence 2, Appli
30	17.2	0.9	25	14	US-10-098-263B-21714	Sequence 21714, A
31	17.2	0.9	32	12	US-10-276-372-24	Sequence 24, Appl
32	17.2	0.9	32	12	US-10-335-394-36	Sequence 26, Appl
33	17.2	0.9	33	9	US-09-812-133-3	Sequence 3, Appli
34	17.2	0.9	35	13	US-10-189-886-2	Sequence 2, Appli
35	17.2	0.9	37	13	US-10-027-632-75809	Sequence 75809, A
36	17.2	0.9	38	10	US-09-960-428-1	Sequence 1, Appli
37	17.2	0.9	38	11	US-09-730-289B-2369	Sequence 2369, Ap
38	17.2	0.9	38	11	US-09-776-474-1355	Sequence 1355, Ap
39	17.2	0.9	38	13	US-10-045-815-13	Sequence 13, Appl
40	17.2	0.9	39	12	US-10-186-185-10	Sequence 10, Appl
41	17.2	0.9	40	11	US-09-998-425-45	Sequence 45, Appl
42	17.2	0.9	40	11	US-09-997-977-45	Sequence 45, Appl
43	17	0.9	25	14	US-10-098-263B-33219	Sequence 33219, A
44	17	0.9	25	14	US-10-098-263B-45507	Sequence 45507, A
45	17	0.9	25	14	US-10-098-263B-59260	Sequence 59260, A

ALIGNMENTS

RESULT 1

US-09-780-929-53
; Sequence 53, Application US/09780929
; Patent No. US20020151693A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc
; APPLICANT: Breaker, Ronald
; APPLICANT: Beigelman, Leo
; TITLE OF INVENTION: Nucleic Acid Catalysts with Endonuclease Activity
; FILE REFERENCE: MBH00-884-H (500/001)
; CURRENT APPLICATION NUMBER: US/09/780,929
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 60/181,360
; PRIOR FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 40
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-780-929-53

Query Match 1.1%; Score 20.4; DB 10; Length 40;
Best Local Similarity 47.4%; Pred. No. 1.1e+05;
Matches 18; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 883 GAATTCGGTTTGGGAATGGATTTTAAAGGAGGGGTATG 920

Db 3 GCAUUGCGUUGAAGAACUGGAAGUUGAAGUGGGCAUG 40

RESULT 2

US-09-152-059-56/C
; Sequence 56, Application US/09152059
; Patent No. US20020068708A1
; GENERAL INFORMATION:
; APPLICANT: WENGEL, JESPER

Query Match 1.0%; Score 18.6; DB 14; Length 25;

US-09-816-467-13

; Sequence 13, Application US/09816467
; Publication No. US20030004121A1

; GENERAL INFORMATION:

; APPLICANT: COEN, LAURENT

; APPLICANT: PINZOLAS, ROSARIO OSTA

; APPLICANT: BRULST, PHILIPPE

; TITLE OF INVENTION: HYBRID PROTEINS THAT MIGRATE RETROGRADELY AND

; FILE REFERENCE: 03495.0174-01000

; CURRENT APPLICATION NUMBER: US/09/816,467

; PRIOR FILING DATE: 2001-03-26

; PRIOR APPLICATION NUMBER: 60/055,615

; PRIOR FILING DATE: 1997-08-14

; PRIOR APPLICATION NUMBER: 60/065,236

; PRIOR FILING DATE: 1997-11-13

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 13

; LENGTH: 40

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Primer

US-09-816-467-13

Query Match 1.0%; Score 18.2; DB 11; Length 40;

Best Local Similarity 74.2%; Pred. No. 4.3e+05;

Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1376 GCAGTCTGTTGTTATAGCACATTTCTTTTGA 1406

Db 2 GCAGTCTCGAGTCTAGACCATGCGCTTTTGA 32

RESULT 11

US-09-801-274-1689/c

; Sequence 1689, Application US/09801274

; Patent No. US20020032319A1

; GENERAL INFORMATION:

; APPLICANT: Cargill, Michele

; APPLICANT: Ireland, James S.

; APPLICANT: Lander, Eric S.

; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS

; FILE REFERENCE: 2825.2009-001

; CURRENT APPLICATION NUMBER: US/09/801,274

; CURRENT FILING DATE: 2001-03-07

; PRIOR APPLICATION NUMBER: US 60/187,510

; PRIOR FILING DATE: 2000-03-07

; PRIOR APPLICATION NUMBER: US 60/206,129

; PRIOR FILING DATE: 2000-05-22

; NUMBER OF SEQ ID NOS: 1802

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1689

; LENGTH: 31

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-801-274-1689

Query Match 1.0%; Score 17.8; DB 9; Length 31;

Best Local Similarity 75.9%; Pred. No. 4.7e+05;

Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1400 TTTTGAAGAATTTCAATAGATTGCA 1428

Db 29 TTTTGAAGACAARTCAGGTTTATGCA 1

RESULT 12

US-10-292-896-25/c

; Sequence 25, Application US/10292896

; Publication No. US20030186850A1

; GENERAL INFORMATION:

; APPLICANT: HASSAN, Helle

US-09-877-478-2673

; APPLICANT: REIS, Celso A.

; APPLICANT: BENNETT, Eric P.

; APPLICANT: CLAUSEN, Henrik

; TITLE OF INVENTION: METHODS TO IDENTIFY AGENTS MODULATING FUNCTIONS OF POLYPEPTIDE GAI

; TITLE OF INVENTION: TRANSFERASES, PHARMACEUTICAL COMPOSITIONS COMPRISING SUCH AGENTS

; FILE REFERENCE: 4305/1H154-US3

; CURRENT APPLICATION NUMBER: US/10/292,896

; CURRENT FILING DATE: 2003-03-17

; PRIOR APPLICATION NUMBER: US 60/425,204

; PRIOR FILING DATE: 2002-11-08

; PRIOR APPLICATION NUMBER: PCT/DK01/00328

; PRIOR FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: US 60/203,331

; PRIOR FILING DATE: 2000-05-11

; NUMBER OF SEQ ID NOS: 130

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 25

; LENGTH: 35

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: primer EBHC75D

US-10-292-896-25

Query Match 1.0%; Score 17.8; DB 12; Length 35;

Best Local Similarity 75.9%; Pred. No. 5.1e+05;

Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 998 GAAGCTGGTCTTCTCTAGAGACTTCG 1026

Db 30 GATGAAGGCTCTTCTTTTAAAGATTCCG 2

RESULT 13

US-09-877-478-2673/c

; Sequence 2673, Application US/09877478

; Publication No. US20030068301A1

; GENERAL INFORMATION:

; APPLICANT: Ribozyme Pharmaceuticals, Inc.

; APPLICANT: Draper, Kenneth

; APPLICANT: Blatt, Larry

; APPLICANT: McSwiggen, Jim

; APPLICANT: Morrissey, Dave

; TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication

; FILE REFERENCE: MBH00-845-H (400/029)

; CURRENT APPLICATION NUMBER: US/09/877,478

; CURRENT FILING DATE: 2001-12-31

; PRIOR APPLICATION NUMBER: US 07/882,712

; PRIOR FILING DATE: 1992-05-14

; PRIOR APPLICATION NUMBER: US 09/531,025

; PRIOR FILING DATE: 2000-03-20

; PRIOR APPLICATION NUMBER: US 09/636,385

; PRIOR FILING DATE: 2000-08-09

; PRIOR APPLICATION NUMBER: US 09/696,347

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 08/193,627

; PRIOR FILING DATE: 1994-02-07

; PRIOR APPLICATION NUMBER: US 08/433,993

; PRIOR FILING DATE: 1995-05-04

; PRIOR APPLICATION NUMBER: US 08/434,504

; PRIOR FILING DATE: 1995-05-04

; PRIOR APPLICATION NUMBER: US 09/436,430

; PRIOR FILING DATE: 1999-11-08

; NUMBER OF SEQ ID NOS: 6586

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 2673

; LENGTH: 38

; TYPE: RNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid

US-09-877-478-2673

Best Local Similarity 75.9%; Pred. No. 5.5e+05;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1753 AAATAAGTACCGTTAAACTTAATCTAAC 1781
|||||
Db 11 AAAGTATATACCGTAAACTTGTCTGAC 39

RESULT 18
US-09-805-761-28/c
; Sequence 28, Application US/09805761
; Patent No. US20020165174A1
; GENERAL INFORMATION:
; APPLICANT: Masood, Rizwan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANTISENSE
; TITLE OF INVENTION: VEGF OLIGONUCLEOTIDES
; FILE REFERENCE: 21327-701CON2
; CURRENT FILING DATE: 2001-03-13
; CURRENT APPLICATION NUMBER: US/09/805,761
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: PCT/US01/00019
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 09/487,023
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: US 09/016,541
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: US 60/037,004
; PRIOR FILING DATE: 1997-01-31
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: VEGF antisense oligonucleotide
US-09-805-761-28

Query Match 0.9%; Score 17.6; DB 10; Length 33;
Best Local Similarity 71.9%; Pred. No. 5.6e+05;
Matches 23; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1498 TCGTGAAGATCGCCAAAGCGAAGCTGTTTACCC 1529
|||||
Db 32 TCGAGTACATCTTCAAGCCATCTCTGTGTGCC 1

RESULT 19
US-10-083-246A-162
; Sequence 162, Application US/10083246A
; Publication No. US20030152936A1
; GENERAL INFORMATION:
; APPLICANT: Athena Diagnostics
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENETIC ANALYSIS OF POLYCYSTIC KIDNEY
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 1133/2002
; CURRENT APPLICATION NUMBER: US/10/083,246A
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(25)
; OTHER INFORMATION: Synthetic primer
US-10-083-246A-162

Query Match 0.9%; Score 17.4; DB 12; Length 25;

Best Local Similarity 94.7%; Pred. No. 5.3e+05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 711 TGACTAAATTTTCTTCAT 729
|||||
Db 7 TGACTAAATTTTCTTCCT 25

RESULT 20
US-09-861-696-46
; Sequence 46, Application US/09861696
; Patent No. US2002007053A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOUPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
; FILE REFERENCE: 11899.0175.CNUS04 MOBT:175-4
; CURRENT APPLICATION NUMBER: US/09/861,696
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/137,440
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 08/833,485
; PRIOR FILING DATE: 1997-04-07
; PRIOR APPLICATION NUMBER: US 08/306,063
; PRIOR FILING DATE: 1994-09-13
; PRIOR APPLICATION NUMBER: US 07/749,611
; PRIOR FILING DATE: 1991-08-28
; PRIOR APPLICATION NUMBER: US 07/576,537
; PRIOR FILING DATE: 1990-08-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 35
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-861-696-46

Query Match 0.9%; Score 17.4; DB 9; Length 35;
Best Local Similarity 77.8%; Pred. No. 6.5e+05;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1026 GTATTCAACATTTGATCTTCAGAA 1052
|||||
Db 2 GAATTCAAACTTCAGATCTTGAGATA 28

RESULT 21
US-09-464-099A-46
; Sequence 46, Application US/09464099A
; Patent No. US20020168680A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOUPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
; FILE REFERENCE: 11899.0175.CNUS01 MOBT:175-2
; CURRENT APPLICATION NUMBER: US/09/464,099A
; CURRENT FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: US 09/137,440
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 08/833,485
; PRIOR FILING DATE: 1997-04-07
; PRIOR APPLICATION NUMBER: US 08/306,063
; PRIOR FILING DATE: 1994-09-13
; PRIOR APPLICATION NUMBER: US 07/749,611
; PRIOR FILING DATE: 1991-08-28
; PRIOR APPLICATION NUMBER: US 07/576,537
; PRIOR FILING DATE: 1990-08-31

```
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 35
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-464-099A-46

Query Match          0.9%; Score 17.4; DB 10; Length 35;
Best Local Similarity 77.8%; Pred. No. 6.5e+05;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1026 GTATTCAAACTTGGATCTTCAGAAA 1052
    ||||| ||||| ||||| ||||| |||||
DB 2 GAATTCAAACTTCAGGATCTTGAGATA 28
    ||||| ||||| ||||| ||||| |||||

RESULT 22
US-10-157-305A-89
; Sequence 89, Application US/10157305A
; Publication No. US20030166099A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Surber, Mark W.
; APPLICANT: Neil Berkley
; APPLICANT: Anca M. Segall
; APPLICANT: Robert Klepper
; TITLE OF INVENTION: MINICELL COMPRISING MEMBRANE PROTEINS
; FILE REFERENCE: MPEX.008DV1
; CURRENT APPLICATION NUMBER: US/10/157,305A
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cloning primer
US-10-157-305A-89

Query Match          0.9%; Score 17.4; DB 12; Length 40;
Best Local Similarity 68.6%; Pred. No. 7.1e+05;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 301 CTGTGGCTTGCAATGCTTATTGTTTGTGATGACG 335
    ||||| ||||| ||||| ||||| |||||
DB 5 CGGTACCTTATCAATGGTGGTGGTGGTGGTGGTGGT 39
    ||||| ||||| ||||| ||||| |||||

RESULT 23
US-10-157-391-89
; Sequence 89, Application US/10157391
; Publication No. US20030166279A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Neil Berkley
; TITLE OF INVENTION: MINICELL-BASED TRANSFECTION
; FILE REFERENCE: MPEX.008DV14
; CURRENT APPLICATION NUMBER: US/10/157,391
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25

US-10-157-391-89
; Sequence 89, Application US/10157305A
; Publication No. US20030190601A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Neil Berkley
; APPLICANT: Surber, Mark W.
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Segall, Anca M.
; APPLICANT: Berkley, Neil
; TITLE OF INVENTION: MINICELL-PRODUCING PARENT CELLS
; FILE REFERENCE: MPEX.008DV23
; CURRENT APPLICATION NUMBER: US/10/157,215A
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24

US-10-157-215A-89
; Sequence 89, Application US/10157215A
; Publication No. US20030190749A1
; GENERAL INFORMATION:
; APPLICANT: Surber, Mark W.
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Segall, Anca M.
; APPLICANT: Berkley, Neil
; TITLE OF INVENTION: MINICELL-PRODUCING PARENT CELLS
; FILE REFERENCE: MPEX.008DV23
; CURRENT APPLICATION NUMBER: US/10/157,215A
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24

US-10-157-096-89
; Sequence 89, Application US/10157096
; Publication No. US20030190601A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Neil Berkley
; APPLICANT: Surber, Mark W.
; TITLE OF INVENTION: TARGET DISPLAY ON MINICELLS
; FILE REFERENCE: MPEX.008DV12
; CURRENT APPLICATION NUMBER: US/10/157,096
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,51
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cloning primer
US-10-157-096-89

Query Match          0.9%; Score 17.4; DB 12; Length 40;
Best Local Similarity 68.6%; Pred. No. 7.1e+05;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 301 CTGTGGCTTGCAATGCTTATTGTTTGTGATGACG 335
    ||||| ||||| ||||| ||||| |||||
DB 5 CGGTACCTTATCAATGGTGGTGGTGGTGGTGGTGGT 39
    ||||| ||||| ||||| ||||| |||||

RESULT 25
US-10-157-215A-89
; Sequence 89, Application US/10157215A
; Publication No. US20030190749A1
; GENERAL INFORMATION:
; APPLICANT: Surber, Mark W.
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Segall, Anca M.
; APPLICANT: Berkley, Neil
; TITLE OF INVENTION: MINICELL-PRODUCING PARENT CELLS
; FILE REFERENCE: MPEX.008DV23
; CURRENT APPLICATION NUMBER: US/10/157,215A
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
```

; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cloning primer
US-10-157-215A-89

Query Match 0.9%; Score 17.4; DB 12; Length 40;
Best Local Similarity 68.6%; Pred. No. 7.1e+05;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 301 CTGTGGCTTGGCAATGCTTATTGTTTGTATGACG 335
DB 5 CGGTACCTTATCAATGGTGTATGATGACG 39

RESULT 26
US-10-157-302-89
; Sequence 89, Application US/10157302
; Publication No. US20030190683A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Surber, Mark W.
; TITLE OF INVENTION: MINICELL-BASED RATIONAL DRUG DESIGN
; FILE REFERENCE: MPEX.008DV17
; CURRENT APPLICATION NUMBER: US/10/157,302
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cloning primer
US-10-157-302-89

Query Match 0.9%; Score 17.4; DB 12; Length 40;
Best Local Similarity 68.6%; Pred. No. 7.1e+05;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 301 CTGTGGCTTGGCAATGCTTATTGTTTGTATGACG 335
DB 5 CGGTACCTTATCAATGGTGTATGATGACG 39

RESULT 27
US-10-154-951B-89
; Sequence 89, Application US/10154951B
; Publication No. US20030194798A1
; GENERAL INFORMATION:
; APPLICANT: Surber, Mark W.
; APPLICANT: Sabbadini, Roger A.
; TITLE OF INVENTION: MINICELL COMPOSITIONS AND METHODS
; FILE REFERENCE: MPEX.008A
; CURRENT APPLICATION NUMBER: US/10/154,951B
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843

; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cloning primer.
US-10-154-951B-89

Query Match 0.9%; Score 17.4; DB 12; Length 40;
Best Local Similarity 68.6%; Pred. No. 7.1e+05;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 301 CTGTGGCTTGGCAATGCTTATTGTTTGTATGACG 335
DB 5 CGGTACCTTATCAATGGTGTATGATGACG 39

RESULT 28
US-10-157-299-89
; Sequence 89, Application US/10157299
; Publication No. US20030194714A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Surber, Mark W.
; TITLE OF INVENTION: MINICELL-BASED TRANSFORMATION
; FILE REFERENCE: MPEX.008DV15
; CURRENT APPLICATION NUMBER: US/10/157,299
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cloning primer
US-10-157-299-89

Query Match 0.9%; Score 17.4; DB 12; Length 40;
Best Local Similarity 68.6%; Pred. No. 7.1e+05;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 301 CTGTGGCTTGGCAATGCTTATTGTTTGTATGACG 335
DB 5 CGGTACCTTATCAATGGTGTATGATGACG 39

RESULT 29
US-10-184-007A-2
; Sequence 2, Application US/10184007A
; Publication No. US20030091546A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Methods of Suppressing Immune Response by Gene Therapy
; FILE REFERENCE: 20263.440
; CURRENT APPLICATION NUMBER: US/10/184,007A
; CURRENT FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 40
; TYPE: DNA
; ORGANISM: artificial

; FEATURE:
; OTHER INFORMATION: reverse primer
US-10-184-007A-2

Query Match 0.9%; Score 17.4; DB 14; Length 40;
Best Local Similarity 77.8%; Pred. No. 7.1e+05;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 665 ACATTCATTTTGGAGATTTCTCTTCA 691
DB 4 ATATCGATTTCAGGCAATTTCTTTTCA 30

RESULT 30

US-10-098-263B-21714/c
; Sequence 21714, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; PRIOR FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 21714
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-098-263B-21714

Query Match 0.9%; Score 17.2; DB 14; Length 25;
Best Local Similarity 86.4%; Pred. No. 6e+05;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1587 CAACAGCTACACTCTATCAGT 1608
DB 25 CAACAGGTACATCTGTATCAGT 4

RESULT 31

US-10-276-372-24/c
; Sequence 24, Application US/10276372
; Publication No. US20030186269A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: SSA-56 kDa Polypeptide and its Fragments and Polynucleotides
; TITLE OF INVENTION: Encoding said Polypeptide and Therapeutic Uses
; FILE REFERENCE: 017753-171
; CURRENT APPLICATION NUMBER: US/10/276,372
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: FR 00/06315
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: PCT/FR 01/00725
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-372-24

Query Match 0.9%; Score 17.2; DB 12; Length 30;
Best Local Similarity 73.3%; Pred. No. 6.7e+05;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 565 CAAGGCTTTTGGAGCCATTTTTCATTCTAA 594
||||||| | ||| ||||| || |||||

DB 30 CAAGGCTGTGGGATCCATGCATGCTTCTCA 1

RESULT 32

US-10-335-394-26/c
; Sequence 26, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Bitagyn, Arva
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/10/335,394
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/646,028
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-10-335-394-26

Query Match 0.9%; Score 17.2; DB 12; Length 32;
Best Local Similarity 73.3%; Pred. No. 7e+05;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 450 GGATTGTTCTTCTATGGGATGGCCGTAG 479
DB 30 GGACCGTCTTCGAGAGGGATGGCCATGG 1

RESULT 33

US-09-812-133-3
; Sequence 3, Application US/09812133
; Patent No. US20020065240A1
; GENERAL INFORMATION:
; APPLICANT: Thomas, Kenneth A., Jr.
; APPLICANT: Kendall, Richard L.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Huckle, William R.
; TITLE OF INVENTION: GENE THERAPY FOR STIMULATION OF
; FILE REFERENCE: 20073P
; CURRENT APPLICATION NUMBER: US/09/812,133
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: PCT/US98/22668
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/063,629
; PRIOR FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: oligonucleotide
US-09-812-133-3

Query Match 0.9%; Score 17.2; DB 9; Length 33;
Best Local Similarity 73.3%; Pred. No. 7.1e+05;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1673 GGGATTCCTCTAGGAACATTATCGTCTCTT 1702
DB 3 GGGATCCAAATATGAACCTTCTGCTCTCTT 32
||||||| | ||| ||||| |||||

RESULT 34
US-10-109-886-2
; Sequence 2, Application US/10109886
; Publication No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; TITLE OF INVENTION: ANTAGONIST TO PPAR
; FILE REFERENCE: TANIGUCHI=6
; CURRENT APPLICATION NUMBER: US/10/109,886
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 35
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially synthesized primer sequence
US-10-109-886-2
Query Match 0.9%; Score 17.2; DB 13; Length 35;
Best Local Similarity 73.3%; Pred. No. 7.4e+05;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 127 CCCATTGAAATCATGCCAGTGTCTCAGG 156
DB 1 CCGCTCAGAAATGTTGGCAGTGGCTCAGG 30
RESULT 35
US-10-027-632-75809
; Sequence 75809, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75809
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-75809
Query Match 0.9%; Score 17.2; DB 13; Length 37;

Best Local Similarity 64.7%; Pred. No. 7.6e+05;
Matches 22; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 66 TTTGACATTGATTTTGAAGAAATTTCTTCAATG 694
DB 2 TTTGATATTGCTCTTARHAAATTTAGTSTAATG 35
RESULT 36
US-09-960-428-1
; Sequence 1, Application US/09960428
; Patent No. US20020115147A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; TITLE OF INVENTION: Method for producing an active heterodimeric AMV-RT in prokaryotic
; FILE REFERENCE: 5272/00/
; CURRENT APPLICATION NUMBER: US/09/960,428
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-960-428-1
Query Match 0.9%; Score 17.2; DB 10; Length 38;
Best Local Similarity 73.3%; Pred. No. 7.7e+05;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1347 CTTTAAATTCATGACCTTTTGGCTACTGTG 1376
DB 6 CTGGAATTCATGACTGTTGCGCTACATCTG 35
RESULT 37
US-09-730-289B-2369/c
; Sequence 2369, Application US/09730289B
; Publication No. US20030050259A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for Treatment of Cardiac Disease
; FILE REFERENCE: MBH00-864-A (400/006)
; CURRENT APPLICATION NUMBER: US/09/730,289B
; CURRENT FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/169,100
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 3897
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2369
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-730-289B-2369
Query Match 0.9%; Score 17.2; DB 11; Length 38;
Best Local Similarity 65.8%; Pred. No. 7.7e+05;
Matches 25; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1226 AGCAAACTATCGAAGAAATTCGTTTATCAGCGACCAT 1263
DB 38 AGCCTACTTTCGGCCTTAACGGCCTCATCAGTTGACCAT 1
RESULT 38
US-09-776-474-1355
; Sequence 1355, Application US/09776474


```
; TITLE OF INVENTION: MMSC1 - An MMAC1 Interacting Protein
; FILE REFERENCE: MMSC1 Gene
; CURRENT APPLICATION NUMBER: US/09/997,977
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 09/233,086
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: US 60/071,861
; PRIOR FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MMSC1 Primers
US-09-997-977-45

Query Match          0.9%; Score 17.2; DB 11; Length 40;
Best Local Similarity 65.8%; Pred. No. 8e+05;
Matches 25; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1603 ATCAGTTTGTAAATGCTTTTGTATTAGCGGCTCCT 1640
      ||||| ||||| ||||| ||||| ||||| |||||
Db 38 ATCAGTATTGTTGGTGACATGTCATAGCTGTTCT 1

RESULT 43
US-10-098-263B-33219/c
; Sequence 33219, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 33219
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-33219

Query Match          0.9%; Score 17; DB 14; Length 25;
Best Local Similarity 80.0%; Pred. No. 6.8e+05;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1682 CTAGGAACCTTTATCGTCTCTTTATA 1706
      ||||| ||||| ||||| ||||| |||||
Db 25 CTAGGATCATTAACGGATCTTTATA 1

RESULT 44
US-10-098-263B-45507
; Sequence 45507, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 45507
; LENGTH: 25
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
US-10-098-263B-45507

Query Match          0.9%; Score 17; DB 14; Length 25;
Best Local Similarity 80.0%; Pred. No. 6.8e+05;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 909 AAGGAGGTATGCCCTTACCTTTAA 933
      ||||| ||||| ||||| ||||| |||||
Db 1 ACGGAGGTATTCGTTTACCTGTAA 25

RESULT 45
US-10-098-263B-59260/c
; Sequence 59260, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 59260
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-59260

Query Match          0.9%; Score 17; DB 14; Length 25;
Best Local Similarity 80.0%; Pred. No. 6.8e+05;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 199 AAAATCTGGAGCGATGCTTTTGT 223
      ||||| ||||| ||||| ||||| |||||
Db 25 ACACCTCTGCAGAGTTGCTTTGT 1

Search completed: October 23, 2003, 12:45:20
Job time : 508 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 02:21:56 ; Search time 500 Seconds
(without alignments)
10063.505 Million cell updates/sec

Title: US-09-868-987-1
Perfect score: 1864
Sequence: 1 atggacttcgcgcatattgtc.....atgacaaattcagataatgc 1864

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 2722572

Minimum DB seq length: 5
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25.4	1.4	88	25	ABX55459 Bovine EST associa
C 2	25.2	1.4	96	22	AAK68003 Human immune/haema
C 3	25	1.3	94	25	ABZ78586 Tumour suppression
C 4	25	1.3	94	25	ABZ09133 Human oligonucleot
C 5	24.4	1.3	90	24	ABA82853 Human protective D
C 6	24.4	1.3	95	20	AAK07251 Genetic suppressor
C 7	24.4	1.3	95	22	AAF93098 Human HIV/tumorige
C 8	24.4	1.3	95	22	AAF94529 Genetic suppressor

9	24.4	1.3	95	24	ABT03434 Human HIV-inhibiti
10	24.2	1.3	65	24	ABN57082 Mouse spliced tran
C 11	24.2	1.3	97	16	AAT26550 Human gene signatu
C 12	24.2	1.3	97	22	AAK68002 Human immune/haema
C 13	24	1.3	64	24	ABZ30164 c-myc-tagged mouse
C 14	24	1.3	65	24	ABZ23420 Candida gene relat
15	24	1.3	78	17	AAT38856 E. coli enterotoxi
16	24	1.3	78	22	AAH28270 Nucleotide sequenc
C 17	23.8	1.3	62	22	AAH28270 Nucleotide sequenc
C 18	23.8	1.3	62	22	AAH28270 Nucleotide sequenc
C 19	23.8	1.3	87	24	ABK76695 Anti-FLK-1 ribozym
C 20	23.8	1.3	89	22	ABK76695 Bacillus lichenifo
21	23.8	1.3	89	22	AAAS1952 Genomic sequence #
C 22	23.6	1.3	60	24	ABN34358 Human spliced tran
C 23	23.6	1.3	65	24	ABZ28946 Candida gene relat
24	23.6	1.3	79	21	AAH11402 Human secreted pro
C 25	23.6	1.3	90	22	AAZ23666 Tetracycline promo
C 26	23.6	1.3	98	24	ABK17244 Coupled ligation a
27	23.6	1.3	98	24	ABK17305 Coupled ligation a
C 28	23.6	1.3	99	22	AAH16979 Protein scaffold m
C 29	23.6	1.3	99	22	AAH16979 Oligonucleotide HF
C 30	23.6	1.3	99	24	ABN87086 Human 10 Fn3 gene
C 31	23.4	1.3	60	24	ABN32824 Human spliced tran
C 32	23.4	1.3	78	21	AAH33034 Arabidopsis thalia
C 33	23.4	1.3	84	22	AAK69024 Human immune/haema
C 34	23.4	1.3	87	16	AAT21076 Human gene signatu
C 35	23.4	1.3	94	21	AAH14189 Human secreted pro
C 36	23.2	1.2	60	24	ABN46009 Human spliced tran
C 37	23.2	1.2	60	24	ABN49485 Human spliced tran
C 38	23.2	1.2	70	21	AAH98467 Nucleotide sequenc
C 39	23.2	1.2	75	21	AAH15734 Human secreted pro
40	23.2	1.2	96	18	AAT65280 Platelet derived g
C 41	23.2	1.2	96	20	AAH87067 Platelet derived g
C 42	23	1.2	76	18	AAT92339 Ribozyme (15) to s
C 43	23	1.2	77	22	ABA74315 Human foetal liver
C 44	23	1.2	77	22	AAK22782 Human brain expres
C 45	23	1.2	77	22	AAK48950 Human bone marrow

ALIGNMENTS

RESULT 1
ABX55459/c
ID ABX55459 standard; CDNA; 88 BP.
XX
AC ABX55459;
XX
DT 26-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #5388.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137160-A1.
XX
PD 26-SEP-2002.
XX
PF 26-OCT-2001; 2001US-0983965.
XX
PR 17-DEC-1998; 98US-113678P.
PR 15-DEC-1999; 99US-0465231.
XX
(BYATT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;

CC (ABZ78294-ABZ79313), which are involved in the molecular pathways of
CC tumour suppression, tumour reversion, apoptosis and/or virus resistance.
CC The sequences are also useful for treatment or prevention of viral,
CC tumour and cell degenerative diseases, and also for diagnosis and
CC prognosis of these diseases.
CC Note: The sequence data for this patent is not represented in the
CC printed specification but is based on sequence information supplied
CC by the European Patent Office.
XX
SQ Sequence 94 BP; 29 A; 17 C; 17 G; 31 T; 0 other;

Query Match 1.3%; Score 25; DB 25; Length 94;
Best Local Similarity 56.8%; Pred. No. 8.2e+03;
Matches 46; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1646 TTAAATTTTGCATTTTATGACCATAGGATTCCTCTAGGAACCTTTATCGTCTCTTTAT 1705
Db 85 TTCAATTTTAAATACACACACAGTAGGAAGTACACTTGGCAGCTGTAAACATGCTCTGTAT 26

QY 1706 ATTGCACCACTCTGTTGTTG 1726
Db 25 ATACACACAGTAGTGCTAG 5

RESULT 4
ABZ09133/C
ID ABZ09133 standard; DNA; 94 BP.
XX
AC ABZ09133;
XX
DT 16-JAN-2003 (first entry)
XX
DE Human oligonucleotide SEQ ID 293.
XX
KW Human; tumour suppressor; virucide; cytostatic; nootropic;
KW neuroprotective; neuroleptic; gene therapy; tumour suppression;
KW tumour reversion; apoptosis; viral resistance; viral infection;
KW cell degeneration; Alzheimer's disease; schizophrenia; cancer; ds.
XX
OS Homo sapiens.
XX
PN FR2822475-A1.
XX
PD 27-SEP-2002.
XX
PF 20-MAR-2002; 2002FR-0003459.
XX
PR 13-FEB-2001; 2001FR-0001925.
XX
PA (MOLE-) MOLECULAR ENGINES LAB SA.
XX
PI Telerman A, Anson R, Tuijnder M, Susini L;
XX
DR WPI; 2003-032204/03.
XX
PT New human nucleic acid, useful for diagnosis, prognosis and treatment,
PT e.g. of tumors, also related vectors, transformed cell, polypeptides
PT and antibodies -
XX
PS Disclosure; Page 82; 189pp; French.
XX
CC The present invention relates to human oligonucleotides
CC (ABZ08941-ABZ09860). The expression of the oligonucleotides is implicated
CC in tumour suppression or reversion, apoptosis and/or viral resistance.
CC The oligonucleotides are useful for preventing and/or treating viral
CC infection, tumour development and cell degeneration (e.g. Alzheimer's
CC disease and schizophrenia), especially cancer.
XX
SQ Sequence 94 BP; 29 A; 17 C; 17 G; 31 T; 0 other;

Query Match 1.3%; Score 25; DB 25; Length 94;
Best Local Similarity 56.8%; Pred. No. 8.2e+03;
Matches 46; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
f. "

QY 1646 TTAAATTTTGCATTTTATGACCATAGGATTCCTCTAGGAACCTTTATCGTCTCTTTAT 1705
Db 85 TTCAATTTTAAATACACACACAGTAGGAAGTACACTTGGCAGCTGTAAACATGCTCTGTAT 26

QY 1706 ATTGCACCACTCTGTTGTTG 1726
Db 25 ATACACACAGTAGTGCTAG 5

RESULT 5
ABA82853/C
ID ABA82853 standard; DNA; 90 BP.
XX
AC ABA82853;
XX
DT 07-FEB-2002 (first entry)
XX
DE Human protective DNA sequence CNI-00746 fragment #18.
XX
KW Human; protective sequence; cell death; cancer; autoimmune disease;
KW neurological disorder; stroke; cytostatic; neuroprotective; gene therapy;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200176457-A2.
XX
PD 18-OCT-2001.
XX
PF 09-APR-2001; 2001WO-US11663.
XX
PR 11-APR-2000; 2000US-0547735.
XX
PA (COGE-) COGENT NEUROSCIENCE INC.
XX
PI Thomas MB, Portbury SD, Puranam K, Katz LC, Lo DC, Barney S;
XX
DR WPI; 2002-025874/03.
DR P-PSDB; ABB44754.
XX
PT New protective sequences and their products, useful for diagnosing and
PT treating diseases involving cell death, including neurological
PT disorders e.g. stroke and for identifying modulators of expression of
PT the protective sequences -
XX
PS Claim 2; Fig 11; 283pp; English.
XX
CC The present invention relates to protective sequence proteins
CC (ABB44624-ABB44830) and their coding sequences (ABA82701-ABA82937).
CC The sequences, when introduced into a cell either predisposed to undergo
CC cell death or in the process of undergoing cell death, prevent, delay or
CC rescue the cell from death, hence, these sequences are named "protective
CC sequences". The sequences are useful for treating and/or ameliorating
CC cancer, autoimmune diseases and neurological disorders e.g. stroke.
CC Further examples of diseases which may be treated by the present
CC invention are given in the specification.
XX
SQ Sequence 90 BP; 29 A; 19 C; 11 G; 31 T; 0 other;

Query Match 1.3%; Score 24.4; DB 24; Length 90;
Best Local Similarity 60.6%; Pred. No. 1.2e+04;
Matches 40; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 713 ACTAAATTTTTCATGCTGTGGATGAATAGACCCCAACATACAGTTGCATATGATG 772
Db 74 AATATGTTGTTTTCATGCTGTGGATGATATAGAAACCCAGCTTAAAGAGTTATGTAGTGG 15

QY 773 AATAG 778
Db 14 AAAAAG 9

RESULT 6

AAAX07251
ID AAX07251 standard; DNA; 95 BP.

XX AC AAX07251;

XX 20-MAR-2003 (updated)

DT 21-MAY-1999 (first entry)

XX Genetic suppressor element CF-313 used to inhibit HIV infection.

XX CF-313; genetic suppressor element; GSE; HIV; infection; AIDS;

KW therapy; protein tyrosine kinase; human; ss.

XX OS Homo sapiens.

XX PN W09854366-A1.

XX PD 03-DEC-1998.

XX 02-JUN-1998; 98WO-US11452.

XX 02-JUN-1997; 97US-0867314.

PR 29-MAY-1998; 98US-0087809.

XX (SUBS-) SUBSIDIARY NO 3 INC.

XX Dayn A, Dunn SJ, Holzmayer TA;

XX WPI; 1999-070224/06.

XX Composition for inhibiting human immunodeficiency virus - containing
PT either fragments of cellular genes that express proteins essential
PT for productive infection, or inhibitors of these cellular gene
PT products

XX Claim 23; Page 60; 85pp; English.

XX This is the nucleotide sequence of genetic suppressor element (GSE)
CC CF-313, which was selected for its ability to prevent productive
CC infection of CEM-s cells by HIV. It has substantial sequence
CC identity with the gene encoding human protein tyrosine phosphatase.
CC The invention relates to compositions and methods for inhibiting
CC HIV infection by down-regulating the expression and/or function of
CC certain human cellular genes. 20 GSEs (see AAX07239-58) are provided
CC that correspond to, or are complementary to, a portion of a
CC cellular gene and which encode intracellular products which are
CC necessary for a productive HIV infection. Some of the selected
CC GSEs function in the sense orientation, possibly as transdominant
CC mutants or RNA decoys. Other GSEs function in the antisense
CC orientation, and may exert their effects as antisense RNA. The
CC GSEs are used to protect cells against HIV infection, either in
CC vitro (e.g. introduced into haematopoietic cells before these are
CC used for engraftment) or in vivo, by direct transduction, e.g.
CC using a gene therapy vector. A method is also provided for
CC identifying additional cellular genes as therapeutic targets for
CC suppressing HIV infection.
CC (Updated on 20-MAR-2003 to correct PR field.)

XX Sequence 95 BP; 28 A; 15 C; 13 G; 39 T; 0 other;

Query Match 1.3%; Score 24.4; DB 20; Length 95;

Best Local Similarity 58.1%; Pred. No. 1.2e+04;

Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 791 AACCATGATTTCTTGAGAGATGCAAAAACCTTGGCGTGTTCCTGGAAGTGTTCCTT 850

Db 14 AAGGATGACCTTTTATACATGAATTAATATGCGATTTCTATTGAAATTTCAACGCT 73

QY 851 TTAGGTTGCGTTGC 864

Db 74 TTTGTTCTTTGGC 87

RESULT 7

AAF93098
ID AAF93098 standard; CDNA; 95 BP.

XX AC AAF93098;

XX 22-MAY-2001 (first entry)

XX Human HIV/tumorigenesis genetic suppressor element, SEQ ID NO:13.

XX HIV infection; tumorigenesis; inhibition; genetic suppressor element;

KW GSE; human cellular gene expression inhibitor; transdominant mutant;

KW RNA decoy; antisense; anti-HIV; anticancer; ss.

XX OS Homo sapiens.

XX PN W0200116323-A2.

XX PD 08-MAR-2001.

XX 01-SEP-2000; 2000WO-US24353.

PR 01-SEP-1999; 99US-0388184.

XX (SUBS-) SUBSIDIARY NO 3 INC.

XX Holzmayer TA, Dunn SJ;

XX WPI; 2001-235108/24.

XX Inhibitory composition for preventing tumorigenesis, comprises a
PT cell-derived nucleic acid molecule, referred to as a genetic suppressor
PT element, a cell-derived protein or an inhibitor of a product of a
PT target gene

XX Example 1; Page 92; 121pp; English.

XX The invention relates to compositions and methods for inhibiting HIV
CC infection and tumorigenesis by downregulating the expression of certain
CC human cellular genes, and/or inhibiting the activity of products encoded
CC by such genes. In particular, the invention relates to a number of human
CC cell-derived nucleic acid molecules which inhibit tumorigenesis or HIV
CC infection in susceptible cells. These nucleic acid molecules, designated
CC genetic suppressor elements (GSEs; AAF93086-AAF93177), correspond to
CC portions of cellular genes or complements thereof. The GSEs in the
CC sense orientation exert their effects as transdominant mutants or
CC RNA decoys, while the GSEs in the antisense orientation act as antisense
CC RNA. The invention also encompasses anti-tumorigenic compositions
CC comprising an isolated cell-derived protein or a mimotope thereof, or an
CC inhibitor of a product of a target gene identified by its ability to
CC inhibit tumour cell formation. The present sequence represents a human
CC GSE described in an exemplification of the invention.

XX Sequence 95 BP; 28 A; 15 C; 13 G; 39 T; 0 other;

Query Match 1.3%; Score 24.4; DB 22; Length 95;

Best Local Similarity 58.1%; Pred. No. 1.2e+04;

Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 791 AACCATGATTTCTTGAGAGGATGCAAAAACCTTGGCGTGTTCCTGGAAGTGTTCCTT 850

Db 14 AAGGATGACCTTTTATACATGAATTAATATGCGATTTCTATTGAAATTTCAACGCT 73

QY 851 TTAGGTTGCGTTGC 864

Db 74 TTTGTTCTTTGGC 87

RESULT 8

AAF94529
ID AAF94529 standard; DNA; 95 BP.

XX

```
AC AAF94529;
XX
XX DT 22-MAY-2001 (first entry)
XX DE
XX Genetic suppressor element DNA sequence CF-313.
XX
XX KW Genetic suppressor element; GSE; human immunodeficiency virus; HIV;
XX infection inhibitor; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200116322-A2.
XX
XX PD 08-MAR-2001.
XX
XX PF 01-SEP-2000; 2000WO-US24262.
XX
XX PR 01-SEP-1999; 99US-0388182.
XX
XX PA (SUBS-) SUBSIDIARY NO 3 INC.
XX
XX PI Holzmayer TA, Dunn SJ;
XX
XX WP; 2001-235107/24.
XX
XX New isolated genetic suppressor element nucleic acid molecule encoding
XX protein such as bone morphogenic protein-1, and double-strand break DNA
XX repair gene protein, for treating human immunodeficiency virus
XX infection.
XX
XX Example 1; Page 91; 106pp; English.
XX
XX Polynucleotide sequences AAF94517 - AAF94536 and AAF94539 - AAF94574
XX represent genetic suppressor element (GSE) nucleic acid molecules which
XX share homology to fragments of genes. When the GSE is operably linked to
XX a regulatory sequence, expression of the GSE in a host cell inhibits
XX infection by human immunodeficiency virus (HIV). The GSE sequences work
XX in gene therapy as human cellular gene product inhibitors or mRNA
XX translation blockers. A composition comprising a protective compound
XX selected from a polypeptide encoded by the GSE sequence, its mimotope,
XX GSE, and an inhibitor of a product of a target gene is useful for
XX protecting host cells from HIV infection by introducing and for treating
XX HIV infection. The GSE sequences are useful for the design of
XX polypeptides or peptides capable of inhibiting HIV infection.
XX
XX Sequence 95 BP; 28 A; 15 C; 13 G; 39 T; 0 other;
XX
XX Query Match 1.3%; Score 24.4; DB 22; Length 95;
XX Best Local Similarity 58.1%; Pred. No. 1.2e+04;
XX Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
XX
XX QY 791 AAGCATGATTCTTGAGAGATGCAAAAACCTTTGGGCTGTTCTCGAAGTGTTCCTT 850
XX
XX Db 14 AAGGATGACTTTTATACAAATGGAATAAATATGCGCATTTCTTATTGAAATTTCAACGCT 73
XX
XX QY 851 TTAGGTTGCGTTCG 864
XX
XX Db 74 TTGTTTCTTTGGC 87
XX
XX Query Match 1.3%; Score 24.4; DB 22; Length 95;
XX Best Local Similarity 58.1%; Pred. No. 1.2e+04;
XX Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
XX
XX QY 791 AAGCATGATTCTTGAGAGATGCAAAAACCTTTGGGCTGTTCTCGAAGTGTTCCTT 850
XX
XX Db 14 AAGGATGACTTTTATACAAATGGAATAAATATGCGCATTTCTTATTGAAATTTCAACGCT 73
XX
XX QY 851 TTAGGTTGCGTTCG 864
XX
XX Db 74 TTGTTTCTTTGGC 87
XX
XX RESULT 9
XX ABT03434
XX ID ABT03434 standard; DNA; 95 BP.
XX
XX AC ABT03434;
XX
XX DT 03-OCT-2002 (first entry)
XX
XX DE Human HIV-inhibiting genetic suppressor element (GSE) 13.
XX
XX KW Human; GSE; genetic suppressor element; ds; HIV inhibition; HIV;
XX HIV biological pathway; HIV inhibitor.
XX
```

```
OS Homo sapiens.
XX
XX PN WO200244722-A2.
XX
XX PD 06-JUN-2002.
XX
XX PF 28-NOV-2001; 2001WO-US44463.
XX
XX PR 28-NOV-2000; 2000US-0724916.
XX
XX PA (SUBS-) SUBSIDIARY NO 3.
XX
XX PI Holzmayer TA, Dunn SJ;
XX
XX WP; 2002-537469/57.
XX
XX Identifying pathways involved in HIV infection, by transferring a
XX fragment library produced from polynucleotide encoding a protein of the
XX pathway into cells and selecting cells having down-regulated latent HIV
XX genes.
XX
XX Example 1; Page 97; 114pp; English.
XX
XX The invention comprises compositions and methods for inhibiting HIV
XX infection by down-regulating expression of certain human cellular genes
XX and/or inhibiting the activity of the products encoded by such genes. The
XX invention particularly relates to genetic suppressor elements (GSEs) -
XX nucleic acid sequences which inhibit HIV infection. The GSEs correspond
XX to portions of cellular genes which encode intracellular products that
XX are necessary for productive HIV infection. The compositions and methods
XX of the invention are useful for identifying a biological pathway involved
XX in HIV infection and for identifying a compound as an HIV inhibitor. The
XX present DNA sequence represents a human GSE of the invention.
XX
XX Sequence 95 BP; 28 A; 15 C; 13 G; 39 T; 0 other;
XX
XX Query Match 1.3%; Score 24.4; DB 24; Length 95;
XX Best Local Similarity 58.1%; Pred. No. 1.2e+04;
XX Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
XX
XX QY 791 AAGCATGATTCTTGAGAGATGCAAAAACCTTTGGGCTGTTCTCGAAGTGTTCCTT 850
XX
XX Db 14 AAGGATGACTTTTATACAAATGGAATAAATATGCGCATTTCTTATTGAAATTTCAACGCT 73
XX
XX QY 851 TTAGGTTGCGTTCG 864
XX
XX Db 74 TTGTTTCTTTGGC 87
XX
XX RESULT 10
XX ABN57082
XX ID ABN57082 standard; DNA; 65 BP.
XX
XX AC ABN57082;
XX
XX DT 15-JUL-2002 (first entry)
XX
XX DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:29830.
XX
XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX OS Mus musculus.
XX
XX PN WO200210449-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 20-JUL-2001; 2001WO-IB01903.
XX
XX PR 28-JUL-2000; 2000US-221607P.
XX
XX PA 02-MAY-2001; 2001US-287724P.
XX
```


(COMP-) COMPUEN INC.
Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
WPI, 2002-257383/30.
New oligonucleotide libraries comprising oligonucleotides which selectively hybridize to mRNAs transcribed from a transcription unit of a genome, useful for detecting tissue-, pathology-, and developmental-specific genes -
Example 1; SEQ ID 29830; 47pp; English.
The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively or quantitatively characterising the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue- and pathology-specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from rats, humans and mice, which are used in the exemplification of the present invention.
N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
Sequence 65 BP; 14 A; 12 C; 17 G; 22 T; 0 other;
Query Match 1.3%; Score 24.2; DB 24; Length 65;
Best Local Similarity 66.0%; Pred. No. 1.1e+04;
Matches 35; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Oy 816 AAAAATTGGCTGTTCTGGAAGTGTTCCTTTTCTTTAGTTCGCTGCTC 868
Db 13 AAAGAATTGGCTGATTTGGTTGAGTTATCCCAATTAGATGCTTCTGCTGC 65
RESULT 11
AAT26550/c
ID AAT26550 standard; cDNA to mRNA; 97 BP.
XX AAT26550;
XX 10-OCT-1996 (first entry)
XX Human gene signature HUMGS08795.
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency; human; cloning; mapping; non-biased library; diagnosis; detection; cell typing; abnormal cell function; ss.
XX Homo sapiens.
XX WO9514772-A1.
XX 01-JUN-1995.
XX 11-NOV-1994; 94WO-JP01916.
XX 12-NOV-1993; 93JP-0355504.
XX

(MATS/) MATSUBARA K.
(OKUB/) OKUBO K.
Matsubara K, Okubo K;
WPI, 1995-206931/27.
Identifying gene signatures in 3'-directed human cDNA library - e.g. for diagnosis of abnormal cell function, by preparing cDNA that reflects relative abundance of corresp. mRNA in specific human tissues
Claim 1; Page 2114; 2245pp; Japanese.
A single-stranded DNA (or its complementary strand or the corresp. double-stranded DNA) which comprises one of the 7837 "GS" sequences given in AAT19001-T26837 and which is able to hybridise to part of human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were obtained from 3'-directed cDNA libraries prepared from various human tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-untranslated sequence is unique to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived. The appearance frequency of a given GS in a cDNA library can be determined (esp. using primers and probes derived from the GS sequences) as a means of diagnosing abnormal cell function or for recognising different cell types.
Sequence 97 BP; 35 A; 8 C; 12 G; 41 T; 1 other;
Query Match 1.3%; Score 24.2; DB 16; Length 97;
Best Local Similarity 62.3%; Pred. No. 1.4e+04;
Matches 38; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Oy 1395 ATTTCTTTTGAAGAAATTCAAATAGATTTGCAAGCCATTGGTCTTTAATGACTGTAT 1454
Db 75 ATTTATTGCAGAAAAATATACAAAGATATTTTACAAAACAATCAATAAATATGATGCAT 16
Oy 1455 T 1455
Db 15 T 15
RESULT 12
AAK68002/c
ID AAK68002 standard; DNA; 97 BP.
XX AAK68002;
XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:22814.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-MAR-2000; 2000US-0198123.
XX

PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220363.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-02559678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 22814; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK61921. (i) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (i)
CC proteins and polynucleotides may be used in the prevention, diagnosis and

CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of a PCR
CC primer used in the method of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

XX SQ Sequence 65 BP; 26 A; 12 C; 8 G; 19 T; 0 other;

Query Match 1.3%; Score 24; DB 24; Length 65;
Best Local Similarity 60.9%; Pred. No. 1.3e+04;
Matches 39; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1684 AGGAACCTTTATCGTCTCTTTATATGACCACCTCTGTTGTTTATGTCGTAAGA 1743
Db 65 ATGGATTTAGCCTCTCATCATATTTTAACGCGTTTGATGTTGATAAGGATTATCAATA 6

QY 1744 AAT 1747
Db 5 GAGT 2

RESULT 15
AAT36856
ID AAT36856 standard; DNA; 78 BP.

XX AC AAT36856;
XX DT 14-MAY-1997 (first entry)
XX DE E. coli enterotoxin secretion signal STII variant.
XX KW Secretion signal; enterotoxin; STII; Escherichia coli; recombinant;
XX PT translation initiation; variant; mutant; reduced translation strength;
XX KW bacterial expression; vector; ds.
XX OS Escherichia coli.
XX OS Synthetic.

XX PN WO9627016-A1.
XX PD 06-SEP-1996.
XX PF 27-FEB-1996; 96WO-US02858.
XX PR 01-MAR-1995; 95US-0398615.
XX PA (GETH) GENENTECH INC.
XX PI Simmons LC, Yansura DG;
XX WPI; 1996-412782/41.

XX PT Optimising secretion of heterologous polypeptide(s) in cells - using
XX PT altered translation initiation regions with reduced translational
XX PT strength compared to wild types
XX PS Claim 4; Page 24; 49pp; English.
XX CC Secretion of heterologous polypeptides of interest in a cell, can
XX CC be optimised by using nucleic acid variants of a translation initiation
XX CC region (TIR), especially signal sequences of the E. coli enterotoxin
XX CC STII. The altered signal sequences result in reduced translational
XX CC strength but high levels of expression of the protein of interest
XX CC and maximum secretion.

XX SQ Sequence 78 BP; 21 A; 14 C; 12 G; 30 T; 1 other;

Query Match 1.3%; Score 24; DB 17; Length 78;
Best Local Similarity 64.3%; Pred. No. 1.4e+04;
Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1659 TTATTATGACCATAGGATTTCTTCTAGGAACCTTTATCGTCTCTTTATATGACCA 1714
Db 11 TGAAGAAAGAACATAGCGTTTCTTCTGTCATCTATGTTCTTTCTATTGTCTACA 66

RESULT 16
AAH28270
ID AAH28270 standard; DNA; 78 BP.
XX AC AAH28270;
XX DT 05-SEP-2001 (first entry)
XX DE Nucleotide sequence of a STII signal sequence variant.
XX KW Protein secretion; signal sequence; STII; ds.
XX OS Synthetic.
XX PN US6242177-B1.
XX PD 05-JUN-2001.
XX PF 01-MAR-1995; 95US-0397303.
XX PR 01-MAR-1995; 95US-0397303.
XX PA (GETH) GENENTECH INC.
XX PI Simmons LC, Yansura DG;
XX WPI; 2001-424299/45.

XX PT Optimizing secretion of a heterologous polypeptide of interest in a
XX PT cell comprises comparing the expression levels of the polypeptide under
XX PT the control of a set of nucleic acid variants of a translation
XX PT initiation region -
XX PS Claim 4; Columns 26; 35pp; English.
XX CC The specification describes a method for the optimising secretion of
XX CC a heterologous polypeptide of interest in a cell. The method comprises
XX CC comparing the expression levels of the polypeptide under the control
XX CC of a set of translation initiation region nucleic acid variants, and
XX CC determining the optimal translation strength for production of mature
XX CC polypeptide, where the strength is less than that of the wild-type
XX CC translation initiation region. The method is useful for increasing or
XX CC optimising the levels of polypeptide secretion. AAH28263-74 represent
XX CC STII signal sequence variants, which are used in the method of the
XX CC invention.

XX SQ Sequence 78 BP; 21 A; 14 C; 12 G; 30 T; 1 other;

Query Match 1.3%; Score 24; DB 22; Length 78;
Best Local Similarity 64.3%; Pred. No. 1.4e+04;
Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1659 TTATTATGACCATAGGATTTCTTCTAGGAACCTTTATCGTCTCTTTATATGACCA 1714
Db 11 TGAAGAAAGAACATAGCGTTTCTTCTGTCATCTATGTTCTTTCTATTGTCTACA 66

RESULT 17
AAS09129
ID AAS09129 standard; DNA; 62 BP.
XX AC AAS09129;
XX DT 24-OCT-2001 (first entry)

```
XX Anti-FLK-1 ribozyme oligonucleotide #1 used to construct plasmid p4562.
DE Human; non-viral ribozyme; tumour growth; metastasis; ReIB; FLK-1;
KW nuclear factor-kappaB2; NFkappaB2; cancer; hyperplastic condition;
KW p4562; ss.
XX Homo sapiens.
OS
XX WO200157061-A1.
PN
XX 09-AUG-2001.
PD
XX 02-FEB-2001; 2001WO-US03406.
PF
XX 04-FEB-2000; 2000US-0180586.
PR
XX (CALP-) CALIFORNIA PACIFIC MEDICAL CENT RES INST.
PA
XX Debs RJ, Kashani-Sabet M;
PI
XX WPI; 2001-488864/53.
DR
XX
XX Identifying function for gene of interest by delivering non-viral
PT ribozyme-encoding polynucleotide into test animal, comparing phenotype
PT of test animal to control and denoting phenotype change as function of
PT gene -
XX
XX Example 1; Page 34; 53pp; English.
PS
XX The present invention relates to a novel method for identifying the
CC function for a gene of interest using ribozymes. The method comprises
CC delivering a non-viral ribozyme-encoding polynucleotide expressing a
CC ribozyme having specificity for the polynucleotide product of the gene
CC of interest into cells of a test animal. The non-viral ribozyme-encoding
CC polynucleotide is useful for treating a disease in an animal, where the
CC polynucleotide sequence is delivered in a non-viral vector. The method
CC is useful for preventing tumour growth of metastasis in a patient by
CC reducing the activity of at least one protein subunit such as ReI, ReIB,
CC nuclear factor-kappaB2 (NFkappaB2), p50 or p65 of NF-kappaB, where the
CC the protein subunit or by delivering a ribozyme specific for an RNA
CC encoding the protein subunit. The method of the invention is useful for
CC treating cancer and hyperplastic conditions. Repeated administration of
CC the polynucleotide encoding a ribozyme is possible without generating an
CC immune response against the vector delivery system, as the
CC the polynucleotide is delivered non-virally. The plasmid vector containing
CC the polynucleotide sequence confers both long term expression of the
CC polynucleotide and the ability to repeatedly reexpress the
CC polynucleotide. The present sequence for anti-FLK-1 ribozyme
CC oligonucleotide #1 is annealed to oligonucleotide #2 (AAS09130) to
CC construct plasmid p4562.
XX
XX Sequence 62 BP; 19 A; 11 C; 14 G; 18 T; 0 other;
SQ
Query Match 1.3%; Score 23.8; DB 22; Length 62;
Best Local Similarity 66.7%; Pred. No. 1.4e+04;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 758 CAGTTGCATATGATGAATAGTTCGTGGGATAAGCATGATTTCTTGAGA 808
DB 10 CGGTACCATTAACTGATGATGTCGTGGAGACGAAACAAAGTTTCTAGAGA 60
RESULT 18
AAS09130/c
ID AAS09130 standard; DNA; 62 BP.
XX
XX AAS09130;
AC
XX 24-OCT-2001 (first entry)
DT
XX Anti-FLK-1 ribozyme oligonucleotide #2 used to construct plasmid p4562.
DE
```

```
XX Human; non-viral ribozyme; tumour growth; metastasis; ReIB; FLK-1;
KW nuclear factor-kappaB2; NFkappaB2; cancer; hyperplastic condition;
KW p4562; ss.
XX Homo sapiens.
OS
XX WO200157061-A1.
PN
XX 09-AUG-2001.
PD
XX 02-FEB-2001; 2001WO-US03406.
PF
XX 04-FEB-2000; 2000US-0180586.
PR
XX (CALP-) CALIFORNIA PACIFIC MEDICAL CENT RES INST.
PA
XX Debs RJ, Kashani-Sabet M;
PI
XX WPI; 2001-488864/53.
DR
XX
XX Identifying function for gene of interest by delivering non-viral
PT ribozyme-encoding polynucleotide into test animal, comparing phenotype
PT of test animal to control and denoting phenotype change as function of
PT gene -
XX
XX Example 1; Page 34-35; 53pp; English.
PS
XX The present invention relates to a novel method for identifying the
CC function for a gene of interest using ribozymes. The method comprises
CC delivering a non-viral ribozyme-encoding polynucleotide expressing a
CC ribozyme having specificity for the polynucleotide product of the gene
CC of interest into cells of a test animal. The non-viral ribozyme-encoding
CC polynucleotide is useful for treating a disease in an animal, where the
CC polynucleotide sequence is delivered in a non-viral vector. The method
CC is useful for preventing tumour growth of metastasis in a patient by
CC reducing the activity of at least one protein subunit such as ReI, ReIB,
CC nuclear factor-kappaB2 (NFkappaB2), p50 or p65 of NF-kappaB, where the
CC the protein subunit or by delivering a ribozyme specific for an RNA
CC encoding the protein subunit. The method of the invention is useful for
CC treating cancer and hyperplastic conditions. Repeated administration of
CC the polynucleotide encoding a ribozyme is possible without generating an
CC immune response against the vector delivery system, as the
CC the polynucleotide is delivered non-virally. The plasmid vector containing
CC the polynucleotide sequence confers both long term expression of the
CC polynucleotide and the ability to repeatedly reexpress the
CC polynucleotide. The present sequence for anti-FLK-1 ribozyme
CC oligonucleotide #2 is annealed to oligonucleotide #1 (AAS09129) to
CC construct plasmid p4562.
XX
XX Sequence 62 BP; 18 A; 14 C; 11 G; 19 T; 0 other;
SQ
Query Match 1.3%; Score 23.8; DB 22; Length 62;
Best Local Similarity 66.7%; Pred. No. 1.4e+04;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 758 CAGTTGCATATGATGAATAGTTCGTGGGATAAGCATGATTTCTTGAGA 808
DB 53 CGGTACCATTAACTGATGATGTCGTGGAGACGAAACAAAGTTTCTAGAGA 3
RESULT 19
ABK76695
ID ABK76695 standard; DNA; 87 BP.
XX
XX ABK76695;
AC
XX 13-AUG-2002 (first entry)
DT
XX Bacillus licheniformis genomic sequence tag (GST) #3986.
DE
XX Differential gene expression; genomic sequenced tag; GST;
KW
```

KW altered culture condition; environmental stress;
KW physiological provocation; ds.

OS Bacillus licheniformis.

PN WO200229113-A2.

PD 11-APR-2002.

XX 05-OCT-2001; 2001WO-US31437.

XX 06-OCT-2000; 2000US-0680598.

PR 27-MAR-2001; 2001US-279526P.

XX (NOVO) NOVOZYMES BIOTECH INC.

PA (NOVO) NOVOZYMES AS.

XX Berka R, Clausen IG;

XX WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array

PS Claim 4; SEQ ID NO 3986; 200pp; English.

CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC a relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive
CC follow-up characterisation is unnecessary, when one spot on an array
CC equals one gene or one open reading frame, since sequence information is
CC available. This sequence represents a genomic sequence tag (GST) used in
CC the method of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 87 BP; 15 A; 18 C; 25 G; 29 T; 0 other;

Query Match 1.3%; Score 23.8; DB 24; Length 87;
Best Local Similarity 59.7%; Pred. No. 1.7e+04;
Matches 40; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 331 GAGCGTATATTAGATTGGAGCGGTCATCGCTCGGAGCTGTTCTTCTGAATCTTTT 390

Db 1 GAGCGGAGGTACGGACACGGAAGCGGTTTGCTTAATTGTTGTTCTGTTCTATTCTCTT 60

QY 391 GCTTATC 397

Db 61 AATCATC 67

RESULT 20

AAS41939/C

ID AAS41939 standard; DNA; 89 BP.

XX AAS41939;

DT 17-DEC-2001 (first entry)

XX Genomic sequence #255 encoding novel human enzyme polypeptide.

XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ds.

XX Homo sapiens.

XX WO200155301-A2..

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01239.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226688.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

KW anti arthritic; nephrotropic; anticoagulant; ds.
XX OS Homo sapiens.
XX WO200155301-A2.
PN
PD
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01239.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-02549678.

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Batash SC, Ruben SM;
 XX WPI; 2001-465566/50.
 XX Novel polypeptides and polynucleotides useful for diagnosing,
 PT preventing, treating neural, immune system, muscular, reproductive,
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
 PT diseases -
 XX Disclosure; SEQ ID No 2078; 1180pp; English.
 XX The present invention relates to the isolation of novel human enzyme
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AAS41695-AAS42192 represent DNA sequences encoding for the novel human
 CC enzyme polypeptides of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 89 BP; 11 A; 23 C; 22 G; 33 T; 0 other;
 Query Match 1.3%; Score 23.8; DB 22; Length 89;
 Best Local Similarity 55.4%; Pred. No. 1.7e+04;
 Matches 46; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
 QY 841 TCTTTTCTTTTAGTTGGTTGCTCTCGGTTGGAGCTCGAATTCGTTTGGGAAT 900
 DB 4 TTTTTTTCTTTTGGACAGAGTCTCACTCTGTTGCCAGGCTGGAGTGGTGGATC 63
 QY 901 GGATTTTAAAGGAGGATGTCCT 923
 DB 64 TTGGCTCACTGAGGCTCTGCTT 86
 RESULT 22
 ABN34358/c
 ID ABN34358 standard; DNA; 60 BP.
 XX
 AC ABN34358;
 XX
 DT 15-JUL-2002 (first entry).
 XX
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:7106.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 XX 20-JUL-2001; 2001WO-IB01903.
 PF
 XX 28-JUL-2000; 2000US-221607P.
 PR
 XX 02-MAY-2001; 2001US-287724P.
 XX

PA (COMP-) COMPUGEN INC.
 XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX WPI; 2002-257383/30.
 XX
 XX New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -
 XX Example 1; SEQ ID 7106; 47pp; English.
 XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridising selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 60 BP; 18 A; 15 C; 13 G; 14 T; 0 other;
 Query Match 1.3%; Score 23.6; DB 24; Length 60;
 Best Local Similarity 76.3%; Pred. No. 1.6e+04;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 447 CTGGGATTGTTCTTCTGCTATGGGATGGCGCTAGATGCA 484
 DB 60 CTGTACTGCTCTGGCTATGGAAGGGCGCTAGATCCA 23
 RESULT 23
 ABZ28946
 ID ABZ28946 standard; DNA; 65 BP.
 XX
 AC ABZ28946;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE Candida gene related tetracyclin promoter PCR primer SEQ ID NO 3029.
 XX
 KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
 KW signal transduction; DNA replication; cell division; growth;
 KW proliferation; Candida albicans; fungicide; antifungal; PCR; primer; ss.
 XX
 OS Candida albicans.
 XX
 PN WO200253728-A2.
 XX
 PD 11-JUL-2002.
 XX
 XX 26-DEC-2001; 2001WO-US49486.
 PF
 XX 29-DEC-2000; 2000US-259128P.
 PR
 XX 20-FEB-2001; 2001US-0792024.
 PR

CC ligation reaction mixture. The reaction mixture is subjected to a cycle
CC of ligation, where adjacently hybridising complementary probes are
CC ligated to one another to form a ligation product. The mixture is then
CC combined with a primer set including a primer comprising a reporter
CC group, and a polymerase to make an amplification mixture. The target
CC sequences can then be detected via the reporter group. The method is
CC useful for identifying splice variants. Sequences ABK17220-ABK17313
CC represent probes and PCR primers used in the method of the invention.
XX

XX Sequence 98 BP; 23 A; 15 C; 20 G; 40 T; 0 other;
SQ

Query Watch 1.3%; Score 23.6; DB 24; Length 98;
Best Local Similarity 58.6%; Pred. No. 2e+04;
Matches 41; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1156 TGTGAGAAACAGCGCTAGATTCTCTACGGAAACTCTAAACGAAACGCAAAATTTTGGTC 1215
DB 77 TGTGACGAAACAGCCATGTATCCATATACGGCAATCAATAAGGAACCTTATGACCTAAAGC 18
QY 1216 AAAGGTAAGC 1225
DB 17 AAAGGTAAGC 8

RESULT 27
ABK17305
ID ABK17305 standard; DNA; 98 BP.
XX
AC ABK17305;
XX
DT 26-MAR-2002 (first entry)
XX
DE Coupled ligation and amplification oligonucleotide #86.
XX
KW Human; coupled ligation and amplification reaction; probe; PCR primer;
XX splice variant; ss.
OS Homo sapiens.
XX
PN WO2001192579-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17329.
XX
PR 10-MAY-2000; 2000US-0584905.
PR 28-NOV-2000; 2000US-0724755.
XX
PA (PEKE) PE CORP NY.
XX
PI Wenz H, ⁶ Schroth GP;
XX
PWPI; 2002-114362/15.
XX
XX
PT Detecting target nucleic acids for identifying splice variants in a
PT target nucleic acid sequence, comprises utilising coupled-ligation and
PT amplification -
XX
XX Disclosure; Page 117; 119pp; English.
XX
XX The invention relates to detecting target sequences in a sample by
XX utilising a coupled ligation and amplification reaction. The method
XX involves combining the sample with a probe set for each target sequence,
XX the probe set comprising a probe comprising a target-specific portion
XX (TSP) and a 5' primer-specific portion or TSP alone, and another probe
XX comprising a TSP and a 3' primer-specific portion, where the probes in
XX each set are suitable for ligation together when hybridised adjacent to
XX one another on a complementary target sequence. At least one probe in
XX each probe set further comprises an addressable support-specific portion
XX located between the primer-specific portion and the TSP to form a
XX ligation reaction mixture. The reaction mixture is subjected to a cycle
XX of ligation, where adjacently hybridising complementary probes are
XX ligated to one another to form a ligation product. The mixture is then

CC combined with a primer set including a primer comprising a reporter
CC group, and a polymerase to make an amplification mixture. The target
CC sequences can then be detected via the reporter group. The method is
CC useful for identifying splice variants. Sequences ABK17220-ABK17313
CC represent probes and PCR primers used in the method of the invention.
XX
SQ Sequence 98 BP; 40 A; 20 C; 15 G; 23 T; 0 other;

Query Match 1.3%; Score 23.6; DB 24; Length 98;
Best Local Similarity 58.6%; Pred. No. 2e+04;
Matches 41; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1156 TGTCAAGAACAGCGCTAGATTCTCTACGGAAGCTCTTAAACGAAAGCAAAATTTGGTC 1215
DB TTTTCTTCCTGTGAAGAGGCTATCCATATACGCAATCAATTAAGGAAGCTTATGACCTAAAGC 81

QY 1216 AAGGTTAAGC 1225
DB TTTTCTTCCTGTGAAGAGGCTATCCATATACGCAATCAATTAAGGAAGCTTATGACCTAAAGC 81

RESULT 28
AAAG4962/c
ID AAA64962 standard; DNA; 99 BP.
XX
AC AAA64962;
XX
DT 06-NOV-2000 (first entry)
XX
DE Protein scaffold molecule encoding sequence HFNLDETOP.
XX
KW Protein scaffold; antibody mimic; fibronectin; protein folding;
KW protein thermostability; ds.
XX
OS Homo sapiens.
XX
PN WO200034784-A1.
XX
PD 15-JUN-2000.
XX
PF 09-DEC-1999; 99WO-US29317.
XX
PR 10-DEC-1998; 98US-0111737.
XX
PA (PHYL-) PHYLLOS INC.
XX
PI Lipovsek D;
XX
DR WPI; 2000-423512/36.
XX
PT Antibody mimics with an ability to bind compound not bound by natural
PT fibronectin useful for evolving novel compound-binding species
PT comprises a fibronectin type III domain having at least one randomized
PT loop -
XX
PS Disclosure; Page 21; 56pp; English.
XX
CC The present sequence, designated HFNLDETOP, encodes part of an antibody
CC mimic based on the fibronectin type III domain. It comprises one of
CC three loops which are designed to be produced randomly, imitating the
CC sequence of the complementarity-determining regions (CDRs) of natural
CC antibody variable regions. The mimics produced using sequences
CC AAA64958-A64970 are useful as they are more stable than naturally
CC occurring proteins, even in conditions which normally lead to their loss
CC of structure. They also exhibit better folding properties than antibody
CC fragments, and can be expressed and folded in prokaryotic, eukaryotic
CC and mammalian systems. They are also able to undergo multiple rounds of
CC selection ('affinity maturation'), involving RNA-protein fusion
CC technology, phage display and yeast display systems.
XX
SQ Sequence 99 BP; 33 A; 28 C; 13 G; 25 T; 0 other;

Query Match 1.3%; Score 23.6; DB 21; Length 99;

Best Local Similarity 58.6%; Pred. No. 2e+04;
Matches 41; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 625 TTTCTTCCTAGATACAGGCGCTATTAAAGGGTTTCCTTGACATTGATTTTAGGAATTTT 684
DB TTTCTTCCTGTGAAGAGGCTATCCATATACGCAATCAATTAAGGAAGCTTATGACCTAAAGC 29

QY 685 CTCCTCAATG 694
DB TTTCTTCCTGTGAAGAGGCTATCCATATACGCAATCAATTAAGGAAGCTTATGACCTAAAGC 29

RESULT 29
AAD16979/c
ID AAD16979 standard; DNA; 99 BP.
XX
AC AAD16979;
XX
DT 29-NOV-2001 (first entry)
XX
DE Oligonucleotide HFNLDETOP.
XX
KW Scaffold protein; antibody mimic; fibronectin type III domain;
KW randomized loop; randomized beta-sheet; diagnostic purpose;
KW protein designing; ss.
XX
OS Unidentified.
XX
PN WO200164942-A1.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06414.
XX
PR 29-FEB-2000; 2000US-0515260.
XX
PA (PHYL-) PHYLLOS INC.
XX
PI Lipovsek D, Wagner RW, Kuimelis RG;
XX
DR WPI; 2001-557782/62.
XX
PT Fibronectin scaffold protein array for obtaining a protein/compound
PT which binds to a compound/protein, comprises a fibronectin type III
PT domain having a randomized loop, a randomized beta-sheet or their
PT combination -
XX
PS Disclosure; Page 23; 67pp; English.
XX
CC The present invention relates to an array of proteins (antibody mimics)
CC comprising a fibronectin type III domain having a randomized loop, a
CC randomized beta-sheet, or their combination, and has the capacity to
CC bind to a compound that is not bound by a corresponding naturally-
CC occurring fibronectin, immobilised onto a solid support. The antibody
CC mimics is useful for detecting a compound preferably a protein, in a
CC biological sample. It is also useful to detect one or more different
CC analytes simultaneously in a sample. Hence is useful for diagnostic
CC purposes. It is also useful for the purpose of designing proteins
CC capable of binding to virtually any compound of interest. The present
CC sequence is an oligonucleotide HFNLDETOP used in an exemplification
CC of the invention.
XX
SQ Sequence 99 BP; 33 A; 28 C; 13 G; 25 T; 0 other;

Query Match 1.3%; Score 23.6; DB 22; Length 99;
Best Local Similarity 58.6%; Pred. No. 2e+04;
Matches 41; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 625 TTTCTTCCTAGATACAGGCGCTATTAAAGGGTTTCCTTGACATTGATTTTAGGAATTTT 684
DB TTTCTTCCTGTGAAGAGGCTATCCATATACGCAATCAATTAAGGAAGCTTATGACCTAAAGC 29

QY 685 CTCCTCAATG 694


```
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 118 TTTAAAGTCCCAATTCAGAAATCATGCCAGTGTCTCAGGGAATTTACC 166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TTTACCCCTTCAAGTTTCAAGTTCATCTCACTGTCTCAGAGAGGTTTCC 50

RESULT 32
AAC53034/C
ID AAC53034 standard; DNA; 78 BP.
XX
AC AAC53034;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 73179.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-0301439.
XX
25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
```

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159395.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 23.4; DB 21; Length 78;
Best Local Similarity 53.4%; Pred. No. 2e+04;
Matches 39; Conservative 4; Mismatches 30; Indels 0; Gaps 0;
QY 1204 AAAATTTTGGCTCAAGTAAAGCAGCAAACTATCGAAGAAAATCGGTTATCAGCGGACCAT 1263
DB 77 AATATTTTTCAGAGAGAAAGAGACTTTATATAATTAAGTAAAGAAAGATTGA 18
QY 1264 CGGCGCTTTTAGGA 1276

Db 17 GAGRAATWTAKGA 5
RESULT 33
AAK69024/c
ID AAK69024 standard; DNA; 84 BP.
XX
AC AAK69024;
XX
DT 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23836.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231141.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-483426/52.
 DR
 XX
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PT
 XX
 PS Disclosure; SEQ ID NO 23836; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent
 CC diagnosis and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 84 BP; 27 A; 20 C; 23 G; 14 T; 0 other;

Query Match 1.3%; Score 23.4; DB 22; Length 84;
 Best Local Similarity 55.6%; Pred. No. 2.1e+04;
 Matches 45; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 843 TTTTCTTTTAGTTCGTTGCTCTCGGGTTTGGAGCTTGGATTCGGTTTGGGAATGG 902
 |||||
 Db 84 TTTTCTTTTAGTTCGTTGCTCTCGGGTTTGGAGCTTGGATTCGGTTTGGGAATGG 902
 |||||

QY 903 ATTTTAAAGGAGGATGCCT 923
 |||||
 Db 24 AGCTCACTGCAAGCTCTGCCT 4
 |||||

RESULT 34
 AAT21076/c
 ID AAT21076 standard; cDNA to mRNA; 87 BP.
 AC AAT21076;
 XX
 XX
 DT 16-JUL-1996 (first entry)
 XX
 XX Human gene signature HUMGS02361.
 XX
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 human; cloning; mapping; non-biased library; diagnosis; detection;

DR WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -

XX Example 1; SEQ ID 18757; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 60 BP; 10 A; 10 C; 20 G; 20 T; 0 other;

Query Match 1.2%; Score 23.2; DB 24; Length 60;
Best Local Similarity 70.5%; Pred. No. 2.1e+04;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 604 AGTATTGGCTCAGCACTCTTTCTTCTTAGATACAGGCGCTA 647

DB 11 AGGAATGGCTCAGCTCTTTTGTCTGCTGTTTAGGGGCCA 54

RESULT 37

ABN49485/C

ID ABN49485 standard; DNA; 60 BP.

XX AC ABN49485;

DT 15-JUL-2002 (first entry)

DE Human spliced transcript detection oligonucleotide SEQ ID NO:22233.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

XX Homo sapiens.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-IB01903.

XX 28-JUL-2000; 2000US-221607P.

XX 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX PI

XX DR

DR WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -

XX Example 1; SEQ ID 22233; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 60 BP; 16 A; 13 C; 17 G; 14 T; 0 other;

Query Match 1.2%; Score 23.2; DB 24; Length 60;
Best Local Similarity 77.8%; Pred. No. 2.1e+04;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 995 CAGGAAGCTGGTCTTTCTTCTTAGAGACTTCGCTATT 1030

DB 46⁶CAGGACGCTGCTCTTTGGGCTTGAGACTTCCTTACT 11

RESULT 38

AAA96467

ID AAA96467 standard; DNA; 70 BP.

XX AC AAA96467;

DT 19-FEB-2001 (first entry)

DE Nucleotide sequence of the directional building block S1.

XX petE promoter; chimeric promoter; transgenic plant; MPr1108;

KW plastocyanin gene promoter; ss.

XX Synthetic.

XX WO200056906-A1.

XX 28-SEP-2000.

XX 20-MAR-2000; 2000WO-IB00317.

XX 22-MAR-1999; 99FR-0003635.

XX (MERI-) MERISTEM THERAPEUTICS.

XX Rance I, Gruber V, Theisen M;

XX WPI; 2000-587667/55.

XX DR

XX Chimeric expression promoter for producing dicotyledonous and
PT monocotyledonous transgenic plants comprises a nucleic acid sequence
PT derived from a promoter of the pea plastocyanin gene
XX
PS Claim 24; Page 72; 83pp; English.
XX
CC The present sequence represents the directional building block 51,
CC which is used to construct chimeric promoters of the invention. The
CC specification describes a chimeric expression promoter comprising a
CC petE promoter of the pea plastocyanin gene, or comprising a G box
CC operably or functionally linked upstream of a CAAT box, TATA box and
CC transcription initiation site. The chimeric promoters are used in
CC expression vectors for producing transgenic plants, such as
CC dicotyledonous species, e.g. potato, tobacco, cotton, lettuce, tomato,
CC melon, cucumber, pea, rape, beetroot, or sunflower, and monocotyledonous
CC species, e.g. wheat, barley, oat, rice, or corn.
XX
SQ Sequence 70 BP; 26 A; 18 C; 10 G; 16 T; 0 other;
Query Match 1.2%; Score 23.2; DB 21; Length 70;
Best Local Similarity 61.7%; Pred. No. 2.2e+04;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 1029 TTCAACATTTGGATCTTCAGAAAGATCAAAATCTATTATTAGTGATAAGCTTTAAGCT 1088
DB 6 TTCAACACATACAAATTCAGTAGAGAGAACTCATTACTCTTGAGAAACCTAGAGGAT 65
RESULT 39
AAC15734/c
ID AAC15734 standard; cDNA; 75 BP.
XX
AC AAC15734;
XX
DT 06-OCT-2000 (first entry)
DE Human secreted protein 5' EST, SEQ ID NO: 19809.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW Gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
DR
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 19809; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.

CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 75 BP; 19 A; 7 C; 13 G; 36 T; 0 other;
Query Match 1.2%; Score 23.2; DB 21; Length 75;
Best Local Similarity 61.7%; Pred. No. 2.3e+04;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 1160 AGAAACAGCGCTAGATTTCTCTACGAAACTTAAACGAAACGAAATTTGGTCAAG 1219
DB 70 AAAAAAATACATGATTTCTCTATGGATATGCACACATACCATATACTTAAAGTCAAG 11
RESULT 40
AAT65280
ID AAT65280 standard; RNA; 96 BP.
XX
AC AAT65280;
XX
DT 10-SEP-1997 (first entry)
XX
DE Platelet derived growth factor AB binding ligand.
XX
KW Platelet derived; growth factor; PDGF; binding ligand;
KW identification; SELEX; anti-mitogenic; inhibition; cell;
KW Systematic Evolution of Ligands by Exponential enrichment;
KW epithelial; proliferation; diagnosis; treatment;
KW neointimal lesion; arterial injury; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..96
FT /*tag= a
FT /note= "all pyrimidines are 2'-fluoro modified"
XX
PN WO9638579-A1.
XX
PD 05-DEC-1996.
XX
PF 30-MAY-1996; 96WO-US08014.
XX
PR 20-MAR-1996; 96US-0618693.
PR 02-JUN-1995; 95US-0458423.
PR 02-JUN-1995; 95US-0458424.
PR 05-JUN-1995; 95US-0465591.
PR 05-JUN-1995; 95US-0465594.
PR 07-JUN-1995; 95US-0479725.
PR 07-JUN-1995; 95US-0479783.
XX
XX (NEXS-) NEXSTAR PHARM INC.
XX
XX Gold L, Janjic N, Pagratia N, Ringquist S, Toothman PJ;
XX
XX WPI; 1997-034387/03.
XX
PT Identification of nucleic acid ligands to TGF-beta, PDGF and hKGF -
PT using SELEX, used in the diagnosis and treatment of proliferative
PT disorders
XX
PS Claim 39; Page 140; 209pp; English.
XX
CC The present sequence, a platelet derived growth factor (PDGF)
CC binding ligand, was identified by Systematic Evolution of Ligands
CC by Exponential enrichment (SELEX). Briefly a candidate mixture of
CC nucleic acids was contacted with PDGF, and nucleic acids having an
CC increased affinity to PDGF partitioned from the remainder of the
CC mixture. The partitioned nucleic acids were then amplified to yield
CC a mixture of nucleic acids enriched for sequences with higher


```

PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 23507; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX SQ Sequence 77-BP; 16 A; 22 C; 4 G; 35 T; 0 other;
XX
Query Match 1.2%; Score 23; DB 22; Length 77;
Best Local Similarity 63.6%; Pred. No. 2.6e+04;
Matches 35; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 726 TCATGCTGTGGATGAATAAGACCCCAACATACACAGTTGTCATATGATGAATAAGTT 780
DB 69 TGATGATGATGAAGACTATGAATAAACATGGGAAGATGCTGATGATAAAGAATTT 15

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 05:46:43 ; Search time 132 Seconds
(without alignments)
6232.865 Million cell updates/sec

Title: US-09-868-987-1
Perfect score: 1864
Sequence: 1 atggacttcgcgcatattgtc.....atgacaaattcagataatgc 1864

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 830080

Minimum DB seq length: 5
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/6C COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24.4	1.3	95	4	US-09-087-609-13
2	24	1.3	78	1	US-08-398-617-22
3	24	1.3	78	2	US-08-398-615-22
4	24	1.3	78	3	US-08-397-303-22
5	23.2	1.2	96	1	US-08-479-783A-67
6	23.2	1.2	96	1	US-08-479-725-67
7	23.2	1.2	96	1	US-08-618-693-67
8	23.2	1.2	96	3	US-08-973-124-156
9	23.2	1.2	96	3	US-08-991-743C-67
10	23.2	1.2	96	4	US-09-851-486-67
11	23.2	1.2	96	5	PCT-US96-08014-156
12	22.8	1.2	85	3	US-08-321-887-36
13	22.8	1.2	100	1	US-08-145-705A-29
14	22.6	1.2	69	3	US-08-584-760A-34
15	22.6	1.2	91	4	US-09-328-750A-5
16	22.6	1.2	96	1	US-08-479-783A-61
17	22.6	1.2	96	1	US-08-479-725-61
18	22.6	1.2	96	1	US-08-618-693-61
19	22.6	1.2	96	3	US-08-973-124-150
20	22.6	1.2	96	3	US-08-991-743C-61
21	22.6	1.2	96	4	US-09-851-486-61
22	22.6	1.2	96	5	PCT-US96-08014-150
23	22.6	1.2	100	1	US-08-472-255A-15
24	22.6	1.2	100	1	US-08-479-724A-15
25	22.6	1.2	100	3	US-08-472-256B-15
26	22.6	1.2	100	3	US-08-952-793-15
27	22.6	1.2	100	4	US-09-849-928-15

c	28	22.6	1.2	100	5	PCT-US96-09455A-15	Sequence 15, Appl
	29	22.4	1.2	72	1	US-08-303-275-195	Sequence 195, App
	30	22.4	1.2	90	3	US-08-974-549A-675	Sequence 675, App
c	31	22.4	1.2	90	3	US-08-974-549A-676	Sequence 676, App
	32	22.2	1.2	47	4	US-09-422-978-1231	Sequence 1231, App
c	33	22.2	1.2	84	6	5217864-2	Patent No. 5217864
c	34	22.2	1.2	96	1	US-08-253-877C-16	Sequence 16, Appl
c	35	22.2	1.2	96	2	US-08-452-164A-16	Sequence 16, Appl
c	36	22.2	1.2	96	3	US-08-603-024-15	Sequence 15, Appl
c	37	22.2	1.2	96	4	US-08-450-809-11	Sequence 11, Appl
c	38	22.2	1.2	97	1	US-08-145-939-7	Sequence 7, Appl
c	39	22.2	1.2	97	1	US-08-343-267-7	Sequence 7, Appl
	40	22	1.2	91	3	US-09-084-120-23	Sequence 23, Appl
	41	21.8	1.2	80	4	US-09-308-759A-18	Sequence 18, Appl
c	42	21.8	1.2	87	3	US-09-237-712-3	Sequence 3, Appl
c	43	21.8	1.2	87	4	US-09-394-457C-13	Sequence 13, Appl
c	44	21.8	1.2	87	4	US-09-709-586A-13	Sequence 13, Appl
c	45	21.8	1.2	87	4	US-09-655-104A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-087-609-13
; Sequence 13, Application US/09087609
; Patent No. 6537972
; GENERAL INFORMATION:
; APPLICANT: Holzmayer, Tanya A.
; APPLICANT: Dunn, Stephen J.
; APPLICANT: Dayn, Andrew
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; INHIBITING HUMAN IMMUNODEFICIENCY VIRUS INFECTION BY
; TITLE OF INVENTION: DOWN-REGULATING HUMAN CELLULAR GENES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,609
; FILING DATE: May 29, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 9320-0009-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-087-609-13

Query Match 1.3%; Score 24.4; DB 4; Length 95;
Best Local Similarity 58.1%; Pred. No. 1.6e+03;
Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

[illegible]

Qy 851 TTAGGTTGCGTTGC 864
||| ||| |||
Db 74 TTTGTTTCTTTGGC 87

RESULT 2
US-08-398-617-22
; Sequence 22, Application US/08398617
; Patent No. 5747662
; GENERAL INFORMATION:
; APPLICANT: Simmons, Laura C.
; APPLICANT: Yansura, Daniel G.
; TITLE OF INVENTION: Methods and Compositions for Secretion of Heterologous Protein
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 460 Pointe San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,617
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:

```

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: P889
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168

```

;
; INFORMATION FOR SEQ ID NO: 22:
;
; SEQUENCE CHARACTERISTICS:
;
;   LENGTH: 78 bases
;
;   TYPE: nucleic acid
;
;   STRANDEDNESS: double
;
;   TOPOLOGY: linear
;
US-08-398-617-22

Query Match          1.3%;   Score 24;   DB 1;   Length 78;
Best Local Similarity 64.3%;   Pred. No. 1.9e+03;

```

	Matches	36; Conservative	0; Mismatches	20; Indels	0; Gaps	0;
Qy	1659	TTATTATGACCATAGGAGATTCTTCTTAGGAAC	TTTATCGTCTCTTTATATTGCACCA	1714		
Db	11	TGAAAAAGAACATAGCGTTCTTCTTTCGATCTATGTCGTTTCTTATGTCTACA	66			
RESULT 3						
US-08-398-615-22						
; Sequence 22, Application US/08398615						

Patent No. 5840523
GENERAL INFORMATION:
APPLICANT: Simmons, Laura C.
APPLICANT: Yamsura, Daniel G.
TITLE OF INVENTION: Methods and Compositions for Secretion of Heterologous Protein
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:

REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: P889
TELEPHONE: 415/225-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 bases
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-397-303-22

Query Match 1.3%; Score 24; DB 3; Length 78;
Best Local Similarity 64.3%; Pred. No. 1.9e+03;
Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1659 TTATTATGACCATAGGATCTCTTAGGAACCTTATCGCTCTTTATATTGCACCA 1714
DB 11 TCAAAAAGAACATAGCGTTCTCTTGCACTATGTTCTGTTTTTCTATTGCTACA 66

RESULT 5

US-08-479-783A-67
Sequence 67, Application US/08479783A
Patent No. 5668264
GENERAL INFORMATION:
APPLICANT: NEBOUSA JANJIC
APPLICANT: LARRY GOLD
TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
TITLE OF INVENTION: ACID LIGANDS
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,783A
FILING DATE: 7-JUNE-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McClearn
REGISTRATION NUMBER: 33,960
REFERENCE/DOCKET NUMBER: NEX42-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:

LENGTH: 96 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F
OTHER INFORMATION: modified
US-08-479-783A-67

Query Match 1.2%; Score 23.2; DB 1; Length 96;
Best Local Similarity 44.7%; Pred. No. 3.5e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

QY 737 ATGAATAAGACCCACATACACAGTTGCATATGATGAATAAGTTCGTGGGGATAAAGCAT 796
DB 8 AAGAUAACGCCUACACCGUGGAGGGGAGGAGUUGAUAUCAGUUGGCGCUCUACUCAU 67
QY 737 GATTTCCTTTGAGAGGATGCAAAAAC 821
DB 68 UCGCCUUCGACAGGAGGCUCACAC 92

RESULT 6

US-08-479-725-67
Sequence 67, Application US/08479725
Patent No. 5674685
GENERAL INFORMATION:
APPLICANT: NEBOUSA JANJIC
APPLICANT: LARRY GOLD
TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
TITLE OF INVENTION: ACID LIGANDS
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,725
FILING DATE: 7-JUNE-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McClearn
REGISTRATION NUMBER: 33,960
REFERENCE/DOCKET NUMBER: NEX42-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:

```
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All pyrimidines are 2'-F
; OTHER INFORMATION: modified
US-08-479-725-67

Query Match      1.2%; Score 23.2; DB 1; Length 96;
Best Local Similarity 44.7%; Pred. No. 3.5e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

QY 737 ATGAATAAGACCCCAACATACACAGTTGCATATGATGAATAGTTCTGGGGATAAGCAT 796
Db 8 AAGAAUAAACCGUACACCGUGAGGGGGAAGUUAUCAGUUGGCGCUCACUCAU 67
QY 797 GATTTCCTTGAGAGGATGCAAAAAAC 821
Db 68 UCGCCUUCGACAGGAGGCUCAAC 92

RESULT 7
US-08-618-693-67
; Sequence 67, Application US/08618693
; Patent No. 5723594
; GENERAL INFORMATION:
; APPLICANT: NEBOJSA JANJIC
; APPLICANT: LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
; TITLE OF INVENTION: ACID LIGANDS
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,693
; FILING DATE: 20 MARCH 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,783
; FILING DATE: 7-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX42/CIP
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All pyrimidines are 2'-F
; OTHER INFORMATION: modified
US-08-618-693-67

Query Match      1.2%; Score 23.2; DB 1; Length 96;
Best Local Similarity 44.7%; Pred. No. 3.5e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

QY 737 ATGAATAAGACCCCAACATACACAGTTGCATATGATGAATAGTTCTGGGGATAAGCAT 796
Db 8 AAGAAUAAACCGUACACCGUGAGGGGGAAGUUAUCAGUUGGCGCUCACUCAU 67
QY 797 GATTTCCTTGAGAGGATGCAAAAAAC 821
Db 68 UCGCCUUCGACAGGAGGCUCAAC 92

RESULT 8
US-08-973-124-156
; Sequence 156, Application US/08973124
; Patent No. 6207816
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD et al.
; TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
; TITLE OF INVENTION: LIGANDS TO GROWTH
; TITLE OF INVENTION: FACTORS
; NUMBER OF SEQUENCES: 304
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,124
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08014
; FILING DATE: 30-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,423
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,424
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,594
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,591
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,725
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,783
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/618,693
; FILING DATE: 20-MARCH-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

QY 737 ATGAATAAGACCCCAACATACACAGTTGCATATGATGAATAGTTCTGGGGATAAGCAT 796
Db 8 AAGAAUAAACCGUACACCGUGAGGGGGAAGUUAUCAGUUGGCGCUCACUCAU 67
QY 797 GATTTCCTTGAGAGGATGCAAAAAAC 821
Db 68 UCGCCUUCGACAGGAGGCUCAAC 92

RESULT 8
US-08-973-124-156
; Sequence 156, Application US/08973124
; Patent No. 6207816
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD et al.
; TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
; TITLE OF INVENTION: LIGANDS TO GROWTH
; TITLE OF INVENTION: FACTORS
; NUMBER OF SEQUENCES: 304
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,124
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08014
; FILING DATE: 30-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,423
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,424
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,594
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,591
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,725
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,783
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/618,693
; FILING DATE: 20-MARCH-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F modified
US-08-973-124-156

Query Match 1.2%; Score 23.2; DB 3; Length 96;
Best Local Similarity 44.7%; Pred. No. 3.5e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

737 ATGAATAAGACCCACATACACAGTTGCATATGATGAATAAGTTCTGTTGGGATAAGCAT 796
8 AAGAAUACGCUAACCCUGGAGGNGGGAAGUUAUCAGUUGUGGCGCUCUACAU 67

797 GATTTCCTTGAGAGGATGCAAAAAC 821
68 UGCGCUUCGACAGGAGGCUACAAC 92

RESULT 9

US-08-991-743C-67
Sequence 67, Application US/08991743C
Patent No. 6229002
GENERAL INFORMATION:
APPLICANT: NEBOJSA JANJIC, LARRY GOLD
TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR (PDGF) NUCLEIC
ACID LIGAND COMPLEXES

NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 1745 Shea Center Drive, Suite 330
CITY: Highlands Ranch
STATE: Colorado
COUNTRY: USA
ZIP: 80129

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,743C
FILING DATE: 16-Dec-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,693
FILING DATE: 20-MARCH-1996
APPLICATION NUMBER: 08/479,783
FILING DATE: 7-JUNE-1995
APPLICATION NUMBER: 08/479,725
FILING DATE: 7-JUNE-1995

ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX66
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 268-0066
TELEFAX: (303) 268-0065

INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA

FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F modified
SEQUENCE DESCRIPTION: SEQ ID NO: 67:

US-08-991-743C-67

Query Match 1.2%; Score 23.2; DB 3; Length 96;
Best Local Similarity 44.7%; Pred. No. 3.5e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

737 ATGAATAAGACCCACATACACAGTTGCATATGATGAATAAGTTCTGTTGGGATAAGCAT 796
8 AAGAAUACGCUAACCCUGGAGGNGGGAAGUUAUCAGUUGUGGCGCUCUACAU 67

797 GATTTCCTTGAGAGGATGCAAAAAC 821
68 UGCGCUUCGACAGGAGGCUACAAC 92

RESULT 10

US-09-851-486-67
Sequence 67, Application US/09851486
Patent No. 6582918
GENERAL INFORMATION:

APPLICANT: NEBOJSA JANJIC, LARRY GOLD
TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR (PDGF) NUCLEIC
ACID LIGAND COMPLEXES

NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,486
FILING DATE: 08-May-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/991,743
FILING DATE: 16-DECEMBER-1997
APPLICATION NUMBER: 08/618,693
FILING DATE: 20-MARCH-1996
APPLICATION NUMBER: 08/479,783
FILING DATE: 7-JUNE-1995
APPLICATION NUMBER: 08/479,725
FILING DATE: 7-JUNE-1995

ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX66
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433

INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA

FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F modified
SEQUENCE DESCRIPTION: SEQ ID NO: 67:

US-09-851-486-67
Query Match 1.2%; Score 23.2; DB 4; Length 96;
Best Local Similarity 44.7%; Pred. No. 3.5e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

737 ATGAATAAGACCCACATACACAGTTGCATATGATGAATAAGTTCTGTTGGGATAAGCAT 796
8 AAGAAUACGCUAACCCUGGAGGNGGGAAGUUAUCAGUUGUGGCGCUCUACAU 67

797 GATTTCCTTGAGAGGATGCAAAAAC 821
68 UGCGCUUCGACAGGAGGCUACAAC 92

```
RESULT 11
PCT-US96-08014-156
; Sequence 156, Application PC/TUS9608014
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD; NBOJSA JANJIC; STEVEN RINGQUIST; NIKOS
; TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
; TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR (TGF ), PLATELET-DERIVED
; TITLE OF INVENTION: GROWTH FACTOR (PDGF) AND HUMAN
; TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR (hKGF)
; NUMBER OF SEQUENCES: 304
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,423
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,424
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,594
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,591
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,725
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,793
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/618,693
; FILING DATE: 20-MARCH-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All pyrimidines are 2'-F modified
PCT-US96-08014-156
Query Match 1.2%; Score 23.2; DB 5; Length 96;
Best Local Similarity 44.7%; Pred. No. 3.5e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;
937 ATGATAAGACCCCAACATACACAGTTCATATGATGAATAAGTTCTGGGGATAAAGCAT 796

RESULT 12
US-08-921-887-36
; Sequence 36, Application US/08921887
; Patent No. 6030771
; GENERAL INFORMATION:
; APPLICANT: KHUDYAKOV, YURI E.
; TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
; TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JONES & ASKEW, LLP
; STREET: 191 Peachtree Street, N.W., 37th Floor
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,887
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WARREN, WILLIAM L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 03063-0380
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis virus
US-08-921-887-36
Query Match 1.2%; Score 22.8; DB 3; Length 85;
Best Local Similarity 62.1%; Pred. No. 4.2e+03;
Matches 36; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
868 CGGGTTTGGAGCTCGAATTCGGTTTGGGAATGGATTTTAAAGGAGGGTATGCCTTT 925
26 CGTATATCGAGCAGGGAATGCTAGCTGAACAGTTTAAACAGAGGCTCTGGCTTT 83

RESULT 13
US-08-145-705A-29/c
; Sequence 29, Application US/08145705A
; Patent No. 5489513
; GENERAL INFORMATION:
; APPLICANT: Springer, Wolfgang; Plempel, Manfred;
; APPLICANT: L bberding, Antonius
; TITLE OF INVENTION: SPECIFIC GENE PROBES AND
; TITLE OF INVENTION: PROCESSES FOR THE DIAGNOSTIC
```

;; TITLE OF INVENTION: INVESTIGATION OF CANDIDA
;; TITLE OF INVENTION: ALBICANS
;; NUMBER OF SEQUENCES: 44
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SPRUNG HORN KRAMER & WOODS
;; STREET: 660 White Plains Road
;; CITY: Tarrytown
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10591-5144
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 MB storage
;; COMPUTER: NEC PowerMate 1 Plus
;; OPERATING SYSTEM: DOS
;; SOFTWARE: WordPerfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/145,705A
;; FILING DATE: October 28, 1993
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: German P 42 36 708.5
;; FILING DATE: October 30, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kurt G. Briscoe
;; REGISTRATION NUMBER: 33,141
;; REFERENCE/DOCKET NUMBER: Bayer 8885-KGB
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (914) 332-1700
;; TELEFAX: (914) 332-1844
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 100 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Candida albicans
;; US-08-145-705A-29

Query Match 1.2%; Score 22.8; DB 1; Length 100;
Best Local Similarity 62.1%; Pred. No. 4.6e+03;
Matches 36; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 1347 CTTTAAATTCATGACCTTTGGTACCTGTGCACTCTTGTATAGCACATTTCTTTT 1404
Db 58 CTTTATTCTTCATATTATCATCATCTGCATTATTTATATATCAAGGCTGTTTT 1

RESULT 14
US-08-584-760A-34/c
; Sequence 34, Application US/08584760A
; Patent No. 6290953
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Courtney, Michael G
; APPLICANT: Finnis, Christopher J A
; APPLICANT: Sleep, Darrell
; TITLE OF INVENTION: Medicine
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/584,760A
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/211,860
;; FILING DATE: 15-APR-1994
;; APPLICATION NUMBER: GB 9121815.6
;; FILING DATE: 14-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Naomi Biswas
;; REGISTRATION NUMBER: 38,384
;; REFERENCE/DOCKET NUMBER: 92H853-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610/878-4294
;; TELEFAX: 610/878/4221
;; INFORMATION FOR SEQ ID NO: 34:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 69 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 1..69
;; OTHER INFORMATION: /function= "oligonucleotide 33"
;; US-08-584-760A-34

Query Match 1.2%; Score 22.6; DB 3; Length 69;
Best Local Similarity 58.0%; Pred. No. 4.4e+03;
Matches 40; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 951 TCAGCGATGTTCCTCAATGCGTGCAAGTTGTGCATAAACTACAGGAAGCTGGTCTTT 1010
Db 69 TCACGTGATATGCTAGATTAGTGATCAAGTTGATGGTCAAGACCATTAGCTGTATTC 10
QY 1011 CTTCTAGAG 1019
Db 9 ATGCTAAG 1

RESULT 15
US-09-328-750A-5/c
; Sequence 5, Application US/09328750A
; Patent No. 6410722
; GENERAL INFORMATION:
; APPLICANT: McGill University et al.
; TITLE OF INVENTION: REPLICATION ORIGIN CONSENSUS SEQUENCES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SWABEY OGILVY RENAULT
; STREET: 1981 McGill College Avenue - Suite 1600
; CITY: Montr.al
; STATE: QC
; COUNTRY: Canada
; ZIP: H3A 2Y3
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,750A
; FILING DATE: 09-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,374

```
;
; FILING DATE: 16-DEC-1996
; APPLICATION NUMBER: 60/047,322
; FILING DATE: 21-MAY-1997
; ATTORNEY/AGENT INFORMATION:
;   NAME: C't., France
;   REGISTRATION NUMBER: 4166
;   REFERENCE/DOCKET NUMBER: 1770-162PCT FC/1d
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 514 845-7126
;   TELEFAX: 514 288-8389
;   TELEX: <unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 91 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; OTHER INFORMATION: SEQ ID NO: 5:
US-09-328-750A-5

Query Match          1.2%; Score 22.6; DB 4; Length 91;
Best Local Similarity 13.6%; Pred. No. 5e+03;
Matches 11; Conservative 47; Mismatches 23; Indels 0; Gaps 0;

QY 1591 AGCTACACTCTATCAGTTTGTGTAATGCTTTTGTATAGGGCGCTCTCTCTTTAA 1650
Db 90 AMYKWMTHMMWMMWMTATHSWTWMMWMMYCMWRYTWMCCWMMTMMWHRKGTSTYDW 31

QY 1651 TTTTGCATTATTATGACCAT 1671
Db 30 BMMWMTYTCWBTHMMWYWT 10

RESULT 16
US-08-479-783A-61
; Sequence 61, Application US/08479783A
; Patent No. 5668264
; GENERAL INFORMATION:
;   APPLICANT: NEBOUSA JANJIC
;   APPLICANT: LARRY GOLD
;   TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
;   TITLE OF INVENTION: ACID LIGANDS
;   NUMBER OF SEQUENCES: 90
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Swanson and Bratschun, L.L.C.
;   STREET: 8400 East Prentice Avenue, Suite #200
;   CITY: Denver
;   STATE: Colorado
;   COUNTRY: USA
;   ZIP: 80111
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
;   COMPUTER: IBM compatible
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/479,783A
;   FILING DATE: 7-JUNE-1995
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/714,131
;   FILING DATE: 10-JUNE-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/931,473
;   FILING DATE: 17-AUGUST-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/964,624
;   FILING DATE: 21-OCTOBER-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/117,991
;   FILING DATE: 8-SEPTEMBER-1993
;   PRIOR APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
;   NAME: Diane H. McClearn
;   REGISTRATION NUMBER: 33,960
;   REFERENCE/DOCKET NUMBER: NEX42-2
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (303) 793-3333
;   TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 96 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All pyrimidines are 2'-F
; OTHER INFORMATION: modified
US-08-479-783A-61

Query Match          1.2%; Score 22.6; DB 1; Length 96;
Best Local Similarity 44.7%; Pred. No. 5.1e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

QY 737 ATGATAAGACCCACATACACAGTTGCATATGATGATGAATAGTTCGTGGGATAAGCAT 796
Db 8 AAGAADAACGCCAACCCGAGUAGGGGGGAAGAAGUUGGCGGCUACUCAU 67

QY 797 GATTTCCTTGAGAGATGCAAAAAC 821
Db 68 UGCGCUUCGACAGAGGCGUCACAC 92

RESULT 17
US-08-479-725-61
; Sequence 61, Application US/08479725
; Patent No. 5674685
; GENERAL INFORMATION:
;   APPLICANT: NEBOUSA JANJIC
;   APPLICANT: LARRY GOLD
;   TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
;   TITLE OF INVENTION: ACID LIGANDS
;   NUMBER OF SEQUENCES: 90
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Swanson and Bratschun, L.L.C.
;   STREET: 8400 East Prentice Avenue, Suite #200
;   CITY: Denver
;   STATE: Colorado
;   COUNTRY: USA
;   ZIP: 80111
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
;   COMPUTER: IBM compatible
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/479,725
;   FILING DATE: 7-JUNE-1995
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/714,131
;   FILING DATE: 10-JUNE-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/931,473
;   FILING DATE: 17-AUGUST-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/964,624
;   FILING DATE: 21-OCTOBER-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/117,991
;   FILING DATE: 8-SEPTEMBER-1993
;   PRIOR APPLICATION DATA:
```



```

;
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All pyrimidines are 2'-F modified
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-991-743C-61

Query Match 1.2%; Score 22.6; DB 3; Length 96;
Best Local Similarity 44.7%; Pred. No. 5.1e+03;
Matches 38; Conservative 8; Mismatches 39; Indels 0; Gaps 0;

QY 737 ATGAATAAGACCAACATACACACAGTTGCATATGATGAATAAGTTCGTGGGGATAAAGCAT 796
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 AAGAAUAAGCCUACACGUGAGUGGGGGGAAAGUUGAAUCAGUUGGCGCUCUACUACAU 67
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 797 GATTTCCTTGAGAGGATGCAAAAAAC 821
   :|||||:|||||:|||||
Db 68 UCGCCUUGACAGGAGGCUCACAC 92
   :|||||:|||||:|||||

RESULT 21
US-09-851-486-61
; Sequence 61, Application US/09851486
; Patent No. 6582918
; GENERAL INFORMATION:
; APPLICANT: NEBOJSA JANJIC, LARRY GOLD
; TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR (PDGF) NUCLEIC
; ACID LIGAND COMPLEXES
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,486
; FILING DATE: 08-May-2001

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/991,743
; FILING DATE: 16-DECEMBER-1997
; APPLICATION NUMBER: 08/618,693
; FILING DATE: 20-MARCH-1996
; APPLICATION NUMBER: 08/479,783
; FILING DATE: 7-JUNE-1995
; APPLICATION NUMBER: 08/479,725
; FILING DATE: 7-JUNE-1995

ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX66

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433

INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All pyrimidines are 2'-F modified
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-851-486-61

Query Match 1.2%; Score 22.6; DB 4; Length 96;

```


Query Match 1.28; Score 22.6; DB 1; Length 100;
Best Local Similarity 68.9; Pred. No. 5.2e+03;
Matches 31; Conservative 0; Mismatches 14; Indels

QY 1345 CGCTTTAAATCATGACACCTTTTGGCTACTGTGCGAGCTTGTTGTTAT 1389

Db 59 CGTTCATAATTCAGTACACTTTTCGATCCCGCCGAGCTCTTGTTGTTAT 15

```

RESULT 26
US08952793-15/c
; Software 15 Application US/08952793
; Patent No 6280932
; GENERAL INFORMATION:
;
; APPLICANT: PARMA, et al.
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
; TITLE OF INVENTION: TO LECTINS

```

CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/O9/849,928
FILING DATE: 04-May-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/952,793
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/479,724
FILING DATE: 07-JUNE-1995
APPLICATION NUMBER: 08/472,256
FILING DATE: 07-JUNE-1995
APPLICANT NUMBER: 08/472,255
FILING DATE: 07-JUNE-1995
APPLICATION NUMBER: 08/477,829
FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX40C/PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:

OTHER INFORMATION: All C's are 2'-NH₂ cytosine

OTHER INFORMATION: All U's are 2'-NH₂ uracil

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-09-849-928-15

Query Match 1.2% Score 22.6; DB 4; Length 100
Best Local Similarity 68.9%; Pred.No. 5.2e+03;
Matches 31; Conservative 0; Mismatches 14; Indels

QY 1345 CGGTTTAATTCAGACACTTGGCTACTGTGCAGTGCTTGTTTTAT 1389
DB 59 CGGTCTAATTCAGTACACTTTCGATCCCGCGAGTCTTGTTGTTAT 15

RESULT 28
PCT-US96-09455A-15/c
Sequence 15, Application PC/TUS9609455A

GENERAL INFORMATION:
APPLICANT: PARMA, et al.
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID
LIGANDS TO LECTINS
NUMBER OF SEQUENCES: 390
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS

```
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09455A
; FILING DATE: 05 JUNE 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,724
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/472,256
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/472,255
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/477,829
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX40C/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All C's are 2'-NH2 cytosine
; FEATURE:
; OTHER INFORMATION: All U's are 2'-NH2 uracil
; PCT-US96-09455A-15

Query Match 1.2%; Score 22.6; DB 5; Length 100;
Best Local Similarity 68.9%; Pred. No. 5.2e+03;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1345 CGCTTTAAATCATGACCTTTGGCTACCTGTGCGAGTCTTGTAT 1389
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 CGTCTAAATCAGTACACATTTTCGATTCCCGCCGAGTCTGTGAT 15

RESULT 29
US-08-303-275-195
; Sequence 195, Application US/08303275
; Patent No. 5766598
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Tartaglia, James
; APPLICANT: Cox, William I.
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS RECOMBINANT
; TITLE OF INVENTION: POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 205
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,275
; FILING DATE:

; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09455A
; FILING DATE: 05 JUNE 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,724
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/472,256
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/472,255
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/477,829
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX40C/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All C's are 2'-NH2 cytosine
; FEATURE:
; OTHER INFORMATION: All U's are 2'-NH2 uracil
; PCT-US96-09455A-15

Query Match 1.2%; Score 22.4; DB 1; Length 72;
Best Local Similarity 59.4%; Pred. No. 5.1e+03;
Matches 38; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 537 TTAAAAAATCTGTAGAAAAAGGATATACCAAGGCTTTTGGAGCCATTTTGTAACT 596
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 TTATAAATCTATCTACATAATCTCTAAAAGGTTCTTTAGGTCCCATTCAGTACAACT 66

QY 597 TGAC 600
|||
Db 67 TAAC 70

RESULT 30
US-08-974-549A-675
; Sequence 675, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
```

```

; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 675:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..90
; OTHER INFORMATION: /note= "oligonucleotide 178"
;
; US-08-974-549A-675
;
; Query Match 1..2%; Score 22.4; DB 3; Length 90;
; Best Local Similarity 56.9%; Pred. NO. 5.6e+03;
; Matches 41; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
;
; QY 1517 AACCTGTTTACCCCTATGATGCTTTTGTAGTAAATGATGCCCTTCAAAGACGTTTCAGCCGC 1576
;
; Db 11 ACGGTGTTTACCCAGGAGATGAATTTTGTGTTACCGAGAAACGACGTTCTGTGTG 70
;
; QY 1577 ACGGTAATGACA 1588
;
; Db 71 ACGGGAACCCCA 82
;
; RESULT 31
; US-08-974-549A-676/c
; Sequence 676, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/974,549A
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 676:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..90
; OTHER INFORMATION: /note= "oligonucleotide 177"
;
; US-08-974-549A-676
;
; Query Match 1..2%; Score 22.4; DB 3; Length 90;
; Best Local Similarity 56.9%; Pred. NO. 5.6e+03;
; Matches 41; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
;
; QY 1517 AACCTGTTTACCCCTATGATGCTTTTGTAGTAAATGATGCCCTTCAAAGACGTTTCAGCCGC 1576
;
; Db 90 ACGGTGTTTACCCAGGAGATGAATTTTGTGTTACCGAGAAACGACGTTCTGTGTG 31
;
; QY 1577 ACGGTAATGACA 1588
;
; Db 30 ACGGGAACCCCA 19
;
; RESULT 32

```

```

US-09-422-978-1231
; Sequence 1231, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CPI
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1231
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-22167-79 : polymorphic base C or T
US-09-422-978-1231

```

```

Query Match 1.2%; Score 22.2; DB 4; Length 47;
Best Local Similarity 69.8%; Pred. No. 4.8e+03;
Matches 30; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 691 AATGTTACGGCTTTTCATGACTAAATTTTCTTCATGCTG 733
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 AAAATTAGATCTCTTCCVTAATAAATTAGGCTTCATCTG 47

```

```

RESULT 33
5217864-2/c
; Patent No. 5217864
; APPLICANT: HEINTZ, NATHANIEL; DAILEY, LISA A.; HEINTZ,
; NICHOLAS H.; CADDLE, MARK S.
; TITLE OF INVENTION: REPLICATION INITIATOR PROTEIN COMPLEX
; AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/573,570
; FILING DATE: 27-AUG-1990
; SEQ ID NO:2:
; LENGTH: 84
5217864-2

```

```

Query Match 1.2%; Score 22.2; DB 6; Length 84;
Best Local Similarity 58.2%; Pred. No. 6.1e+03;
Matches 39; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 519 TTTTATGCTCAAGCTTAAATAATCTGTAGAAAAGGATATACCAAGGCTTTTGGAG 578
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 TTTTGTGATCAAACTCAGTAAAGAACTAAATAATAATAATAATAATAATAATAATAATA 8

```

```

QY 579 CCATTTT 585
|||||
Db 7 TCATTTT 1

RESULT 34
US-08-253-877C-16/c
; Sequence 16, Application US/08253877C
; Patent No. 5773001
; GENERAL INFORMATION:
; APPLICANT: Hamann, Philip R.
; APPLICANT: Hamann, Lois
; APPLICANT: Hollander, Irwin

```

```

; APPLICANT: Holcomb, Ryan
; APPLICANT: Hallett, William
; APPLICANT: Tsou, Hwei-Ru
; APPLICANT: Weiss, Martin J.
; TITLE OF INVENTION: Conjugates of Methyltrithio Antitumor
; Agents and Intermediates for Their Synthesis
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/253,877C
; FILING DATE: 03-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 32,368
; TELEPHONE: 201-831-3246
; TELEFAX: 201-831-3305
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-253-877C-16

Query Match 1.2%; Score 22.2; DB 1; Length 96;
Best Local Similarity 77.1%; Pred. No. 6.5e+03;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 407 CTACAGTATTTGGATGCGCCACTCACCTTGTCAGG 441
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 CTACATTATTTGGATGAGACAGGACCTGGACAGG 39

RESULT 35
US-08-452-164A-16/c
; Sequence 16, Application US/08452164A
; Patent No. 5877296
; GENERAL INFORMATION:
; APPLICANT: Hamann, Philip R.
; APPLICANT: Hinman, Lois
; APPLICANT: Hollander, Irwin
; APPLICANT: Holcomb, Ryan
; APPLICANT: Hallett, William
; APPLICANT: Tsou, Hwei-Ru
; APPLICANT: Weiss, Martin J.
; TITLE OF INVENTION: Conjugates of Methyltrithio Antitumor
; Agents and Intermediates for Their Synthesis
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products Corporation
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

US-08-603-024-15

NUMBER OF SEQUENCES: 15

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o Robbins, Berliner & Carson
; STREET: 201 No. 554501th Figueroa Street, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,939
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Spitals, John P.
; REGISTRATION NUMBER: 29,215
; REFERENCE/DOCKET NUMBER: 5727-109C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 977-1001
; TELEFAX: (213) 977-1003
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-145-939-7

Query Match 1.2%; Score 22.2; DB 1; Length 97;
Best Local Similarity 54.2%; Pred. No. 6.5e+03;
Matches 45; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 572 TTGTGGAGCCATTCTTGTATCTTAAGG 654
|||||
DB 88 TTGTGATGACGCTTCTGTATCTATATCATAGGAACCAAGATGATATTTCTTT 29
|||||

QY 632 CTAGATACAGGCGCTATTAAAGG 654
|||||
DB 28 AATGGTCCAGGCATAATCCAGG 6
|||||

RESULT 39
US-08-343-267-7/c
; Sequence 7, Application US/08343267
; Patent No. 5583211
; GENERAL INFORMATION:
; APPLICANT: Coassin, Peter J.
; APPLICANT: Matson, Robert S.
; APPLICANT: Rampal, Jang B.
; TITLE OF INVENTION: Biopolymer Synthesis
; TITLE OF INVENTION: Utilizing Surface
; TITLE OF INVENTION: Activated, Organic
; TITLE OF INVENTION: Polymers
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beckman Instruments, Inc.
; STREET: 2500 Harbor Boulevard
; CITY: Fullerton
; STATE: California
; COUNTRY: USA
; ZIP: 92634
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44
; MEDIUM TYPE: Mb
; COMPUTER: IBM
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,267
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/971,100
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Burgoon, Richard P.
; REGISTRATION NUMBER: 34,787
; REFERENCE/DOCKET NUMBER: 129D-141
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 773-6969
; TELEFAX: (714) 773-7936
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: no
; ANTI-SENSE: no
; US-08-343-267-7

Query Match 1.2%; Score 22.2; DB 1; Length 97;
Best Local Similarity 54.2%; Pred. No. 6.5e+03;
Matches 45; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 572 TTGTGGAGCCATTCTTGTATCTTAAGG 654
|||||
DB 88 TTGTGATGACGCTTCTGTATCTATATCATAGGAACCAAGATGATATTTCTTT 29
|||||

QY 632 CTAGATACAGGCGCTATTAAAGG 654
|||||
DB 28 AATGGTCCAGGCATAATCCAGG 6
|||||

RESULT 40
US-09-084-120-23
; Sequence 23, Application US/09084120
; Patent No. 6251592
; GENERAL INFORMATION:
; APPLICANT: TANG, JianQing
; APPLICANT: MELANCON, Serge B.
; TITLE OF INVENTION: A NOVEL STR MARKER SYSTEM
; TITLE OF INVENTION: FOR DNA FINGERPRINTING
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SWABEY OGILVY RENAULT
; STREET: Suite 1600, 1981 McGill College Avenue
; CITY: Montreal
; STATE: QC
; COUNTRY: Canada
; ZIP: H3A 2Y3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,120
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: COTE, France
; REGISTRATION NUMBER: 37,037
; REFERENCE/DOCKET NUMBER: 13251-4US FC/ld
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 514 845-7126
```



```
; TELEFAX: 514 288-8389
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-084-120-23

Query Match 1.2%; Score 22; DB 3; Length 91;
Best Local Similarity 59.7%; Pred. No. 7.2e+03;
Matches 37; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1103 CGAGCTCTCTCTCTAAATAACGATCATGAGCTGGCGTTATTGTGGGATTGTGTGAGA 1162
DB 6 CGAGACTCTGTCANATATAACAAACAAACTCTGGTCTCTGGCGCTCTCTTGAAA 65

QY 1163 AA 1164
DB 66 TA 67

RESULT 41
US-09-308-759A-18
; Sequence 18, Application US/09308759A
; Patent No. 6391593
; GENERAL INFORMATION:
; APPLICANT: Weston, Anthony
; Assenberg, Rene
; Marsh, Peter
; Mock, Graham A.
; Ray, Trevor D.
; Wharam, Susan D.
; Cardy, Donald L.N.
; TITLE OF INVENTION: Modified Nucleic Acid Probes and Uses Thereof
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESS: PILLSBURY WINTHROP LLP
; STREET: 1600 TYSONS BOULEVARD
; CITY: McLean
; STATE: VA
; COUNTRY: USA
; ZIP: 22102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/308,759A
; FILING DATE: 19-Jul-1999
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-308-759A-18

Query Match 1.2%; Score 21.8; DB 4; Length 80;
Best Local Similarity 65.3%; Pred. No. 7.7e+03;
Matches 32; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 809 GGATGCAAAAACCTTGGCTGTTCTTGGAGTGTTTTTCTTTTAGTT 857
DB 32 GAATGCAAAACACTGAGAGTGTCTTGAGATGCTTCTGTTGATTT 80

RESULT 42
US-09-237-712-3/c

; Sequence 3, Application US/09237712
; Patent No. 6180391
; GENERAL INFORMATION:
; APPLICANT: BROWN, WILLIAM C.
; TITLE OF INVENTION: HIGHLY EFFICIENT CONTROLLED EXPRESSION OF EXOGENOUS
; FILE REFERENCE: A-518
; CURRENT APPLICATION NUMBER: US/09/237,712
; CURRENT FILING DATE: 1999-01-26
; EARLIER APPLICATION NUMBER: 60/072,794
; EARLIER FILING DATE: 1998-01-28
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 87
; TYPE: DNA
; ORGANISM: T4 bacteriophage
US-09-237-712-3

Query Match 1.2%; Score 21.8; DB 3; Length 87;
Best Local Similarity 61.4%; Pred. No. 8e+03;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1652 TTTCATTTATTATGACCATAGGATTTCTTCTAGGAACCTTTATCGTCTCTTTATTT 1708
DB 81 TTATTTCTATTATAACCATATGGATTATTAAAGCAAGCTTCTTTTCGTTTGCTTT 25

RESULT 43
US-09-394-457C-13/c
; Sequence 13, Application US/09394457C
; Patent No. 6440705
; GENERAL INFORMATION:
; APPLICANT: Variagenics, Inc.
; TITLE OF INVENTION: A Method for Analyzing Polynucleotides
; FILE REFERENCE: 246/020
; CURRENT APPLICATION NUMBER: US/09/394,457C
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 13
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hypothetical sequence to demonstrate application...
US-09-394-457C-13

Query Match 1.2%; Score 21.8; DB 4; Length 87;
Best Local Similarity 65.3%; Pred. No. 8e+03;
Matches 32; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1334 AGTCCGATGCGGCTTTTAATTCATGACCTTTGGCTACCTGTGCACTCT 1382
DB 69 AGGCACATATGAATTTTCATTCAGCAGCTTGATGTCGCGGTGAAGTCT 21

RESULT 44
US-09-709-596A-13/c
; Sequence 13, Application US/09709596A
; Patent No. 6458945
; GENERAL INFORMATION:
; APPLICANT: Variagenics, Inc.
; TITLE OF INVENTION: A Method for Analyzing Polynucleotides
; FILE REFERENCE: 258/239
; CURRENT APPLICATION NUMBER: US/09/709,596A
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 13
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Hypothetical sequence to demonstrate application..
US-09-709-596A-13
Query Match      1.2%; Score 21.8; DB 4; Length 87;
Best Local Similarity 65.3%; Pred. No. 8e+03;
Matches 32; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 1334 AGTCCGCGTATGCCCTTTAATTCATGACCTTTTGCTACCTGTGCACTCT 1382
      |||||
Db 69 AGGACATATGAATTTTCATTACAGCAGCTTGATGGTCCCGGTGAAGTCT 21
```

```
RESULT 45
US-09-655-104A-13/c
; Sequence 13, Application US/09655104A
; Patent No. 6500650
; GENERAL INFORMATION:
; APPLICANT: Variagenics, Inc.
; TITLE OF INVENTION: A Method for Identifying Polymorphisms
; FILE REFERENCE: 257/078
; CURRENT APPLICATION NUMBER: US/09/655,104A
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hypothetical sequence to demonstrate application..
US-09-655-104A-13
```

```
Query Match      1.2%; Score 21.8; DB 4; Length 87;
Best Local Similarity 65.3%; Pred. No. 8e+03;
Matches 32; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 1334 AGTCCGCGTATGCCCTTTAATTCATGACCTTTTGCTACCTGTGCACTCT 1382
      |||||
Db 69 AGGACATATGAATTTTCATTACAGCAGCTTGATGGTCCCGGTGAAGTCT 21
```

Search completed: October 23, 2003, 09:22:31
Job time : 133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 04:31:57 ; Search time 6720 Seconds
(without alignments)
11347.551 Million cell updates/sec

Title: US-09-868-987-1

Perfect score: 1864

Sequence: 1 atggacttcgcgatattgtc.....atgacaaattcagataatgc 1864

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 1519770

Minimum DB seq length: 5

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.man:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	25.6	1.4	64	11	BX293717	BX293717 Arabidops
2	25.6	1.4	81	6	BD113070	BD113070 EST and e
3	25.6	1.4	81	6	BD113370	BD113370 EST and e
C 4	25.4	1.4	91	14	MAARNA4AX	M10826 Alfalfa mos
C 5	25	1.3	94	6	AX522623	AX522623 Sequence
6	24.8	1.3	94	3	S79433	S79433 cels [5' un
7	24.8	1.3	99	14	AY144374	AY144374 Emiliana
8	24.4	1.3	81	9	HSU81086	U81086 Human immun
9	24.4	1.3	87	11	BX296209	BX296209 Arabidops
10	24.4	1.3	95	6	AR300761	AR300761 Sequence
11	24.4	1.3	95	6	AX089664	AX089664 Sequence
12	24.4	1.3	95	6	AX090119	AX090119 Sequence
13	24.4	1.3	95	6	AX543215	AX543215 Sequence
14	24.2	1.3	97	8	ATH521008	AJ521008 Arabidops
15	24.2	1.3	100	8	AY200525	AY200525 Arabidops
C 16	24	1.3	65	6	AX486203	AX486203 Sequence
17	24	1.3	78	6	AR005344	AR005344 Sequence
18	24	1.3	78	6	AR059872	AR059872 Sequence
19	24	1.3	78	6	AR156129	AR156129 Sequence
C 20	23.8	1.3	54	8	ATH522698	AJ522698 Arabidops
21	23.8	1.3	87	6	AX435571	AX435571 Sequence
22	23.8	1.3	87	8	AF187253	AF187253 Polytrich
C 23	23.8	1.3	93	8	ATH526349	AJ526349 Arabidops
C 24	23.8	1.3	100	1	LACP1AA	M63184 L.lactis P1
25	23.6	1.3	65	6	AX485729	AX485729 Sequence
C 26	23.6	1.3	72	3	MIAATVR	X03925 Mosquito mi
27	23.6	1.3	79	6	BD035147	BD035147 Sequence
28	23.6	1.3	90	6	AX241109	AX241109 Sequence
C 29	23.6	1.3	98	6	AX326604	AX326604 Sequence
30	23.6	1.3	98	6	AX326665	AX326665 Sequence
C 31	23.4	1.3	88	10	MMGCH3	AF090727 Mus muscu
C 32	23.4	1.3	91	8	YSCWTP041	J01518 Yeast [S.ce
33	23.4	1.3	94	6	BD037934	BD037934 Sequence
34	23.2	1.2	70	6	AX036605	AX036605 Sequence
C 35	23.2	1.2	75	6	BD039479	BD039479 Sequence
36	23.2	1.2	86	8	ATH530983	AJ530983 Arabidops
37	23.2	1.2	91	3	S37388	S37388 Trypanosoma
38	23.2	1.2	96	6	AR140879	AR140879 Sequence
39	23.2	1.2	96	6	AR150829	AR150829 Sequence
40	23.2	1.2	96	6	I65707	I65707 Sequence 67
41	23.2	1.2	96	6	I67939	I67939 Sequence 67
42	23.2	1.2	96	6	I90160	I90160 Sequence 67
C 43	23.2	1.2	98	6	BD115800	BD115800 EST and e
44	23.2	1.2	99	14	AY144376	AY144376 Emiliana
45	23.2	1.2	99	14	AY144377	AY144377 Emiliana

ALIGNMENTS

RESULT 1
BX293717/c
LOCUS
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.23352, sequence tagged site.
ACCESSION BX293717
VERSION BX293717.1 GI:28949549
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,

Murphy,G., Langham,S., Legrys,C., Jones,J.D.G. and Bevan,M.
Unpublished
2 (bases 1 to 64)
Clarke,J.H.
Direct Submission
Submitted (10-MAR-2003) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
AT denotes an activation tag dissociation transposon within a
single line, ET an enhancer trap dissociation transposon, GT a gene
trap dissociation transposon, MT a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. _3 denotes a sequence derived from the 3' end of the
transposon, _5 denotes a sequence derived from the 5' end of the
transposon BBSRC GARNET, ATIS project
On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock
code: DEAD.

FEATURES
source
1..64
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AC018848"
/note="Derived from superpool 16.28 NASC code N40723"
1..64
/standard_name="SM_3.23352"
34 a 3 c 3 g 24 t

STS
BASE COUNT 34 a 3 c 3 g 24 t
ORIGIN

Query Match 1.4%; Score 25.6; DB 11; Length 64;
Best Local Similarity 62.5%; Pred. No. 4.8e+05;
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1020 ACTTCGATTCAACATTTGGATCTTCAGAAAGATCAAAATCTATTTAGTGATAAG 1079
|||||
Db 64 ACTTTTAAATCAATAATGAATTAAATATAATATATATATATATTTTGTGTTAGAAA 5
|||||

QY 1080 CTTT 1083
|||||
Db 4 CTTT 1

RESULT 2
BD113070 81 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION EST and encoded human protein.
ACCESSION BD113070
VERSION BD113070.1 GI:23207888
KEYWORDS JP 2002010789-A/5147.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
EST and encoded human protein
Patent: JP 2002010789-A 5147 15-JAN-2002;
GENSET CORP

OS Homo sapiens (human)
PN JP 2002010789-A/5147
PD 15-JAN-2002
PF 07-AUG-2000 JP 20020780989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1..81
/organism="Homo sapiens (human)".

REFERENCE
AUTHORS
TITLE
JOURNAL
PATENT: JP 2002010789-A 5147 15-JAN-2002;
GENSET CORP

OS Homo sapiens (human)
PN JP 2002010789-A/5147
PD 15-JAN-2002
PF 07-AUG-2000 JP 20020780989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1..81
/organism="Homo sapiens (human)".

FEATURES
source
1..81
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
16 a 11 c 8 g 45 t 1 others

BASE COUNT 16 a 11 c 8 g 45 t 1 others
ORIGIN

Query Match 1.4%; Score 25.6; DB 6; Length 81;
Best Local Similarity 57.5%; Pred. No. 4.7e+05;
Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1419 TAGATTGCAAGCCATTGGTCTTTAAATGACTGTATTTGGGTATTCATTAAACAATACCTT 1478
|||||
Db 1 TAAATTTTGAGCCCAATGTTACAAATTAAGTCTTTCTGAGAGTATTTTATACCCCTGTAT 60
|||||

QY 1479 TGATCATTTTTCATCGTATT 1498
|||||
Db 61 TTTTCTTCTCTTTTCTTTT 80
|||||

RESULT 3
BD113370 81 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION EST and encoded human protein.
ACCESSION BD113370
VERSION BD113370.1 GI:23208188
KEYWORDS JP 2002010789-A/5447.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
EST and encoded human protein
Patent: JP 2002010789-A 5447 15-JAN-2002;
GENSET CORP

OS Homo sapiens (human)
PN JP 2002010789-A/5447
PD 15-JAN-2002
PF 07-AUG-2000 JP 2002080989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1..81
/organism="Homo sapiens (human)".

FEATURES
source
1..81
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
16 a 11 c 8 g 46 t

BASE COUNT 16 a 11 c 8 g 46 t
ORIGIN

Query Match 1.4%; Score 25.6; DB 6; Length 81;
Best Local Similarity 57.5%; Pred. No. 4.7e+05;
Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1419 TAGATTGCAAGCCATTGGTCTTTAAATGACTGTATTTGGGTATTCATTAAACAATACCTT 1478
|||||
Db 1 TAAATTTTGAGCCCAATGTTACAAATTAAGTCTTTCTGAGAGTATTTTATACCCCTGTAT 60
|||||

QY 1479 TGATCATTTTTCATCGTATT 1498
|||||
Db 61 TTTTCTTCTCTTTTCTTTT 80
|||||

Db 85 TTCAATTTTAAATACACACAGTAGGACACTGGGCACCTGTACATGCTCTGTAT 26

Qy 1706 ATGCACCACTCTGTGTG 1726
 ==
 Db 25 ATACACACAAGTATGTGCTAG 5
 ==

RESULT 6
 S79433
 LOCUS
 DEFINITION 94 bp mRNA linear INV 30-NOV-1995
 cels [5' untranslated region] [Dictyostelium discoideum=slime mold,
 mRNA Partial, 94 nt].

ACCESSION S79433
 VERSION S79433.1 GI:1087019
 KEYWORDS
 SOURCE
 ORGANISM Dictyostelium discoideum
 Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 REFERENCE 1 (bases 1 to 94)
 AUTHORS Ramalingam,K., Blume,J.E., Ganguly,K. and Ennis,H.L.
 TITLE AT-rich upstream sequence elements regulate spore
 germination-specific expression of the Dictyostelium discoideum
 cels gene
 JOURNAL Nucleic Acids Res. 23 (15), 3018-3025 (1995)
 MEDLINE 95388521
 PUBMED 7659526
 REMARK GenBank staff at the National Library of Medicine created this
 entry [NCBI gibbsq 170735] from the original journal article.
 This sequence comes from Table 1.

FEATURES
 source
 1..94
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /db_xref="taxon:44689"
 gene
 1..94
 /partial
 . /gene="cels"
 BASE COUNT 24 a 6 c 1 g 63 t
 ORIGIN

Query Match 1.3%; Score 24.8; DB 3; Length 94;
 Best Local Similarity 60.3%; Pred. No. 7e+05,
 Matches 41; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1600 TCTATCAGTTTCTTAATGCTTTTGGTTATGCGCGCTCCTCTGTCTTAAATTTGCATT 1659
 ===
 Db 7 TCTATCAATTTTCAAAATTTTTTTTATACAGTATCTTTTTTTTATTTTAAATTT 66
 ===

Qy 1660 TATTATGA 1667
 ==
 Db 67 TTTTAA 74
 ==

RESULT 7
 AY144374
 LOCUS
 DEFINITION 99 bp DNA linear VRL 07-MAY-2003
 Emiliania huxleyi virus isolate OTU1 major capsid protein (MCP)

ACCESSION AY144374
 VERSION AY144374.1 GI:24429757
 KEYWORDS
 SOURCE
 ORGANISM Emiliania huxleyi virus
 Emiliania huxleyi virus
 Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae.

REFERENCE 1 (bases 1 to 99)
 AUTHORS Schroeder,D.C., Oke,J., Hall,M., Malin,G. and Wilson,W.H.
 TITLE Virus Succession Observed during an Emiliania huxleyi Bloom
 JOURNAL Appl. Environ. Microbiol. 69 (5), 2484-2490 (2003)
 PUBMED 12732512
 REFERENCE 2 (bases 1 to 99)
 AUTHORS Schroeder,D.C., Oke,J., Hall,M., Evans,C., Malin,G. and Wilson,W.H.
 TITLE Direct Submission
 JOURNAL Submitted (26-AUG-2002) Marine Biological Association, Citadel

QY	556	AGGATATACCAAGGCTTTTGAGCCATTTTGAFTTCACTACAG	605
Db	1	ATGTTTACTATGTTGGGAGTATTATGATTCTACTTTGACTACTG	50
RESULT 9			
LOCUS	BX296209	87 bp	DNA linear STS 10-JUN-2003
DEFINITION	Arabidopsis thaliana transposon insertion STS SM_3.35057, sequence tagged site.		
ACCESSION	BX296209		
VERSION	BX296209.1	GI:29170074	
KEYWORDS	STS; STS, sequence tagged site.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1		
AUTHORS	Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 87)		
AUTHORS	Clarke, J.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK		
COMMENT	AT denotes an activation tag dissociation transposon within a single line, Et an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. 3 denotes a sequence derived from the 3' end of the transposon. 5 denotes a sequence derived from the 5' end of the transposon. BSRG GARNET, ATIS project		
On-line seed stock requests:	http://nasc.mott.ac.uk/ NASC stock code: N121768.		
FEATURES			
source	1..87		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/variety="Columbia-0 NASC stock code N1092"		
	/db_xref="taxon:3702"		
	/clone="AC009918"		
	/note="Derived from superpool 26.44 NASC code Unknown"		
STS	1..87		
	/standard_name="SM_3.35057"		
BASE COUNT	54 a 4 c 12 g 17 t		
ORIGIN			
Query Match	1.3%;	Score 24.4;	DB 11; Length 87;
Best Local Similarity	60.6%;	Pred. No. 8.6e+05;	
Matches	40;	Conservative 0;	Mismatches 26; Indels 0; Gaps 0;
QY	1597	AACTCTATCAGTTTGTAAATGCTTTTGTATAGGGCGCTCTGCTTTAAATTTGC	1656
Db	80	AACTTTATTCGTTTTTATTTATTTTTCATTTTTCATTTTCTTAATCTGA	21
QY	1657	ATTAT 1662	
Db	20	TAATAT 15	
RESULT 10			
LOCUS	AR300761	95 bp	DNA linear PAT 12-JUN-2003
DEFINITION	Sequence 13 from patent US 6537972.		
ACCESSION	AR300761		
VERSION	AR300761.1	GI:31688328	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
QY	556	AGGATATACCAAGGCTTTTGAGCCATTTTGAFTTCACTACAG	605
Db	1	ATGTTTACTATGTTGGGAGTATTATGATTCTACTTTGACTACTG	50
RESULT 9			
LOCUS	BX296209/c		
DEFINITION	Arabidopsis thaliana transposon insertion STS SM_3.35057, sequence tagged site.		
ACCESSION	BX296209		
VERSION	BX296209.1	GI:29170074	
KEYWORDS	STS; STS, sequence tagged site.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1		
AUTHORS	Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 87)		
AUTHORS	Clarke, J.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK		
COMMENT	AT denotes an activation tag dissociation transposon within a single line, Et an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. 3 denotes a sequence derived from the 3' end of the transposon. 5 denotes a sequence derived from the 5' end of the transposon. BSRG GARNET, ATIS project		
On-line seed stock requests:	http://nasc.mott.ac.uk/ NASC stock code: N121768.		
FEATURES			
source	1..87		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/variety="Columbia-0 NASC stock code N1092"		
	/db_xref="taxon:3702"		
	/clone="AC009918"		
	/note="Derived from superpool 26.44 NASC code Unknown"		
STS	1..87		
	/standard_name="SM_3.35057"		
BASE COUNT	54 a 4 c 12 g 17 t		
ORIGIN			
Query Match	1.3%;	Score 24.4;	DB 11; Length 87;
Best Local Similarity	60.6%;	Pred. No. 8.6e+05;	
Matches	40;	Conservative 0;	Mismatches 26; Indels 0; Gaps 0;
QY	1597	AACTCTATCAGTTTGTAAATGCTTTTGTATAGGGCGCTCTGCTTTAAATTTGC	1656
Db	80	AACTTTATTCGTTTTTATTTATTTTTCATTTTTCATTTTCTTAATCTGA	21
QY	1657	ATTAT 1662	
Db	20	TAATAT 15	
RESULT 10			
LOCUS	AR300761	95 bp	DNA linear PAT 12-JUN-2003
DEFINITION	Sequence 13 from patent US 6537972.		
ACCESSION	AR300761		
VERSION	AR300761.1	GI:31688328	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
QY	556	AGGATATACCAAGGCTTTTGAGCCATTTTGAFTTCACTACAG	605
Db	1	ATGTTTACTATGTTGGGAGTATTATGATTCTACTTTGACTACTG	50
RESULT 9			
LOCUS	BX296209/c		
DEFINITION	Arabidopsis thaliana transposon insertion STS SM_3.35057, sequence tagged site.		
ACCESSION	BX296209		
VERSION	BX296209.1	GI:29170074	
KEYWORDS	STS; STS, sequence tagged site.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1		
AUTHORS	Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.		
JOURNAL	Unpublished		
REFERENCE			

REFERENCE 1 (bases 1 to 95)
AUTHORS Holzmayer,T.A., Dunn,S.J. and Dayn,A.
TITLE Compositions and methods for inhibiting human immunodeficiency virus infection by down-regulating human cellular genes
JOURNAL Patent: US 6537972-A 13 25-MAR-2003;
FEATURES Location/Qualifiers
source 1..95
BASE COUNT 28 a 15 c 13 g 39 t
ORIGIN
Query Match 1.3%; Score 24.4; DB 6; Length 95;
Best Local Similarity 58.1%; Pred. No. 8.6e+05;
Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 791 AACCATGATTTCTTGAGAGGATGCAAAAACCTTTGGCTGTTTCTGGAAGTGTTCCTT 850
Db 14 AAGCATGACCTTTTATACATGGAATAAATTATGCAATTTCTATTGAAATTTCAACGCT 73
QY 851 TTAGGTTCGTTGC 864
Db 74 TTGTTTCTTTGGC 87
RESULT 11
LOCUS AX089664 95 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 13 from Patent WO0116322.
ACCESSION AX089664
VERSION AX089664.1 GI:13443846
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Holzmayer,T.A. and Dunn,S.J.
TITLE Compositions and methods for inhibiting human immunodeficiency virus infection by down-regulating human cellular genes
JOURNAL Patent: WO 0116322-A 13 08-MAR-2001;
Subsidiary N0. 3, INC. (US)
FEATURES Location/Qualifiers
source 1..95
BASE COUNT 28 a 15 c 13 g 39 t
ORIGIN
Query Match 1.3%; Score 24.4; DB 6; Length 95;
Best Local Similarity 58.1%; Pred. No. 8.6e+05;
Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 791 AACCATGATTTCTTGAGAGGATGCAAAAACCTTTGGCTGTTTCTGGAAGTGTTCCTT 850
Db 14 AAGCATGACCTTTTATACATGGAATAAATTATGCAATTTCTATTGAAATTTCAACGCT 73
QY 851 TTAGGTTCGTTGC 864
Db 74 TTGTTTCTTTGGC 87
RESULT 12
LOCUS AX090119 95 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 13 from Patent WO0116323.
ACCESSION AX090119
VERSION AX090119.1 GI:13444080
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Holzmayer,T.A. and Dunn,S.J.
TITLE Compositions and methods for inhibiting human immunodeficiency virus infection by down-regulating human cellular genes
JOURNAL Patent: WO 0116322-A 13 08-MAR-2001;
Subsidiary N0. 3, INC. (US)
FEATURES Location/Qualifiers
source 1..95
BASE COUNT 28 a 15 c 13 g 39 t
ORIGIN
Query Match 1.3%; Score 24.4; DB 6; Length 95;
Best Local Similarity 58.1%; Pred. No. 8.6e+05;
Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 791 AACCATGATTTCTTGAGAGGATGCAAAAACCTTTGGCTGTTTCTGGAAGTGTTCCTT 850
Db 14 AAGCATGACCTTTTATACATGGAATAAATTATGCAATTTCTATTGAAATTTCAACGCT 73
QY 851 TTAGGTTCGTTGC 864
Db 74 TTGTTTCTTTGGC 87

REFERENCE 1
AUTHORS Holzmayer,T.A. and Dunn,S.J.
TITLE Compositions and methods for inhibiting human immunodeficiency virus infection by down-regulating human cellular genes
JOURNAL Patent: WO 0116323-A 13 08-MAR-2001;
Subsidiary N0. 3, INC. (US)
FEATURES Location/Qualifiers
source 1..95
BASE COUNT 28 a 15 c 13 g 39 t
ORIGIN
Query Match 1.3%; Score 24.4; DB 6; Length 95;
Best Local Similarity 58.1%; Pred. No. 8.6e+05;
Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 791 AACCATGATTTCTTGAGAGGATGCAAAAACCTTTGGCTGTTTCTGGAAGTGTTCCTT 850
Db 14 AAGCATGACCTTTTATACATGGAATAAATTATGCAATTTCTATTGAAATTTCAACGCT 73
QY 851 TTAGGTTCGTTGC 864
Db 74 TTGTTTCTTTGGC 87
RESULT 13
LOCUS AX543215 95 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 13 from Patent WO0244722.
ACCESSION AX543215
VERSION AX543215.1 GI:25276515
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Holzmayer,T.A. and Dunn,S.J.
TITLE Compositions and methods for inhibiting human immunodeficiency virus infection by down-regulating human cellular genes
JOURNAL Patent: WO 0244722-A 13 06-JUN-2002;
Subsidiary N0. 3, INC. (US)
FEATURES Location/Qualifiers
source 1..95
BASE COUNT 28 a 15 c 13 g 39 t
ORIGIN
Query Match 1.3%; Score 24.4; DB 6; Length 95;
Best Local Similarity 58.1%; Pred. No. 8.6e+05;
Matches 43; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 791 AACCATGATTTCTTGAGAGGATGCAAAAACCTTTGGCTGTTTCTGGAAGTGTTCCTT 850
Db 14 AAGCATGACCTTTTATACATGGAATAAATTATGCAATTTCTATTGAAATTTCAACGCT 73
QY 851 TTAGGTTCGTTGC 864
Db 74 TTGTTTCTTTGGC 87
RESULT 14
LOCUS ATH521008 97 bp DNA linear PLN 29-MAR-2003
DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone 050C12.
ACCESSION AJ521008
VERSION AJ521008.1 GI:26789244
KEYWORDS left border; T-DNA flanking sequence.

SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 AUTHORS Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., Lepiniec, L., Caboche, M., and Lecharny, A.
 TITLE T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 97)
 AUTHORS Balzerque, S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
 COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.inbioigen.fr>).

FEATURES
 source
 1..97
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="WassiljewskiJa"
 /db_xref="taxon:3702"
 /clone="050C12"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 misc_feature
 1..97
 /note="T-DNA flanking sequence
 left border"
 26 a 21 c 29 g 21 t

BASE COUNT
 ORIGIN
 Query Match 1.3%; Score 24.2; DB 8; Length 97;
 Best Local Similarity 66.0%; Pred. No. 9.5e+05;
 Matches 35; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 322 TGTATTCATCAGCGTATATTATAGATTGGAGCGTCATCGCTTCGGAGCTG 374
 |||||
 Db 10 TGATTTGATCAGCGTATGATTTGATTAAAGGCGTCAGACGCTAGGATG 62

RESULT 15
 AY200525 100 bp DNA linear PLN 24-JAN-2003
 LOCUS Arabidopsis thaliana sequence flanking Ds3 end of Ds-EnhancerTrap
 DEFINITION Insertion from line EF8745.
 ACCESSION AY200525
 VERSION AY200525.1 GI:27896479
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 100)
 AUTHORS May, B.P., Simorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R., McCombie, W.R., and Martienssen, R.A.
 TITLE Arabidopsis genomic sequences flanking Ds enhancer and gene traps
 JOURNAL In transgenic lines
 REFERENCE 2 (bases 1 to 100)

AUTHORS May, B.P., Simorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R., McCombie, W.R., and Martienssen, R.A.
 TITLE Direct Submission
 JOURNAL Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
 COMMENT <http://genetrap.cshl.org>.
 FEATURES
 Location/Qualifiers
 source
 1..100
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /clone="EF8745.Ds3.07.21.2001.jw03.100"
 /note="transgenic line EF8745
 ecotype: Landsberg"
 29 a 23 c 10 g 38 t

BASE COUNT
 ORIGIN
 Query Match 1.3%; Score 24.2; DB 8; Length 100;
 Best Local Similarity 62.3%; Pred. No. 9.4e+05;
 Matches 38; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1351 AATTATGACCTTTGGTACCTGCTGAGTCTGTTATAGCATTCTTTTGAAGAA 1410
 |||||
 Db 21 AACTTATAAAGTGTGCATATCATGCTCTCTTTTAAATAACCTTTCTTTGGCAGCA 80
 |||||
 QY 1411 A 1411
 |||||
 Db 81 A 81

RESULT 16
 AX486203/c 65 bp DNA linear PAT 16-AUG-2002
 LOCUS AX486203 Sequence 3503 from Patent WO20053728.
 DEFINITION AX486203
 ACCESSION AX486203.1 GI:22320419
 VERSION
 KEYWORDS
 SOURCE Candida albicans
 ORGANISM Candida albicans
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.
 REFERENCE 1
 AUTHORS Roemer, T., Jiang, B., Boone, C., Bussey, H., and Ohlsen, K.L.
 TITLE Gene disruption methodologies for drug target discovery
 JOURNAL Patent: WO 02053728-A 3503 11-JUL-2002;
 Elitra Pharmaceuticals, Inc. (US)
 FEATURES
 Location/Qualifiers
 source
 1..65
 /organism="Candida albicans"
 /mol_type="genomic DNA"
 /db_xref="taxon:5476"
 26 a 12 c 8 g 19 t

BASE COUNT
 ORIGIN
 Query Match 1.3%; Score 24; DB 6; Length 65;
 Best Local Similarity 60.9%; Pred. No. 1.1e+06;
 Matches 39; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1684 AGGAACCTTATCGTCTCTTTATTTATTTGACCACTCTGTTGTTTATGTCGTAAGA 1743
 |||||
 Db 65 ATGGATTTTAGCCTCTCATATATTTAAACGCGTTTGATGTTGATAAGGATTATCAATA 6

QY 1744 AAT 1747
 |||||
 Db 5 GAGT 2

RESULT 17
 AR005344 78 bp DNA linear PAT 04-DEC-1998
 LOCUS AR005344 Sequence 22 from patent US 5747662.
 DEFINITION AR005344
 ACCESSION AR005344.1 GI:3966223
 VERSION

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 78)
 AUTHORS Simmons,L.C. and Yansura,D.G.
 TITLE Methods and compositions for secretion of heterologous polypeptides
 JOURNAL Patent: US 5747662-A 22 05-MAY-1998;
 FEATURES
 Location/Qualifiers
 1..78
 /organism="unknown"
 BASE COUNT 21 a 14 c 12 g 30 t 1 others
 ORIGIN
 Query Match 1.3%; Score 24; DB 6; Length 78;
 Best Local Similarity 64.3%; Pred. No. 1.1e+06;
 Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1659 TTATTATGACCATAGGATCTCTTAGGAACCTTTATCGTCTTTATATGACCA 1714
 Db 11 TGAAGAAAGAACATAGCGTTCTCTTGCACTATGTTGCTTTTCTATTGCTACA 66
 RESULT 18
 AR059872
 LOCUS
 DEFINITION Sequence 22 from patent US 5840523.
 ACCESSION AR059872
 VERSION AR059872.1 GI:5986322
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 78)
 AUTHORS Simmons,L.C. and Yansura,D.G.
 TITLE Methods and compositions for secretion of heterologous polypeptides
 JOURNAL Patent: US 5840523-A 22 24-NOV-1998;
 FEATURES
 Location/Qualifiers
 1..78
 /organism="unknown"
 BASE COUNT 21 a 14 c 12 g 30 t 1 others
 ORIGIN
 Query Match 1.3%; Score 24; DB 6; Length 78;
 Best Local Similarity 64.3%; Pred. No. 1.1e+06;
 Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1659 TTATTATGACCATAGGATCTCTTAGGAACCTTTATCGTCTTTATATGACCA 1714
 Db 11 TGAAGAAAGAACATAGCGTTCTCTTGCACTATGTTGCTTTTCTATTGCTACA 66
 RESULT 19
 AR156129
 LOCUS
 DEFINITION Sequence 22 from patent US 6242177.
 ACCESSION AR156129
 VERSION AR156129.1 GI:15124833
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 78)
 AUTHORS Simmons,L.C. and Yansura,D.G.
 TITLE Methods and compositions for secretion of heterologous polypeptides
 JOURNAL Patent: US 6242177-A 22 05-JUN-2001;
 FEATURES
 Location/Qualifiers
 1..78
 /organism="unknown"
 BASE COUNT 21 a 14 c 12 g 30 t 1 others
 ORIGIN
 Query Match 1.3%; Score 24; DB 6; Length 78;
 Best Local Similarity 64.3%; Pred. No. 1.1e+06;
 Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1659 TTATTATGACCATAGGATCTCTTAGGAACCTTTATCGTCTTTATATGACCA 1714
 Db 11 TGAAGAAAGAACATAGCGTTCTCTTGCACTATGTTGCTTTTCTATTGCTACA 66
 RESULT 19
 AR156129
 LOCUS
 DEFINITION Sequence 22 from patent US 6242177.
 ACCESSION AR156129
 VERSION AR156129.1 GI:15124833
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 78)
 AUTHORS Simmons,L.C. and Yansura,D.G.
 TITLE Methods and compositions for secretion of heterologous polypeptides
 JOURNAL Patent: US 6242177-A 22 05-JUN-2001;
 FEATURES
 Location/Qualifiers
 1..78
 /organism="unknown"
 BASE COUNT 21 a 14 c 12 g 30 t 1 others
 ORIGIN
 Query Match 1.3%; Score 24; DB 6; Length 78;
 Best Local Similarity 64.3%; Pred. No. 1.1e+06;
 Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1659 TTATTATGACCATAGGATCTCTTAGGAACCTTTATCGTCTTTATATGACCA 1714
 Db 11 TGAAGAAAGAACATAGCGTTCTCTTGCACTATGTTGCTTTTCTATTGCTACA 66

Best Local Similarity 64.3%; Pred. No. 1.1e+06;
 Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1659 TTATTATGACCATAGGATCTCTTAGGAACCTTTATCGTCTTTATATGACCA 1714
 Db 11 TGAAGAAAGAACATAGCGTTCTCTTGCACTATGTTGCTTTTCTATTGCTACA 66
 RESULT 20
 ATH522698/c
 LOCUS
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
 297A10.
 ACCESSION AJ522698
 VERSION AJ522698.1 GI:26790934
 KEYWORDS left border; T-DNA flanking sequence.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1
 AUTHORS Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Sanson,F.,
 Chauvin,S., Bechtold,N., Cruaud,C., Derose,R., Pelletier,G.,
 Lepiniec,L., Caboche,M. and Lecharny,A.
 TITLE T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 54)
 AUTHORS Balzergue,S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
 Gaston Cremieux, 91057 Evry cedex, FRANCE
 COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
 plants from INRA (Versailles). The DNA fragment(s) resulting from
 the PCR were directly sequenced from the left or the right border
 to determine the genomic sequence flanking the insertion. T-DNA
 derived sequences were removed. Information to order the
 corresponding mutant line and a link to a database providing a
 graphical display of the insertion site are available at
 http://dbgap.versailles.inra.fr/publiclines/. This sequence has
 been generated in the framework of the French plant genomics
 program 'Genoplante' (http://www.genoplante.com and
 http://genoplante-info.infobiogen.fr).
 FEATURES
 Location/Qualifiers
 1..54
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Wassiljewskija"
 /db_xref="taxon:3702"
 /clone="297A10"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 misc_feature 1..54
 /note="T-DNA flanking sequence
 left border"
 BASE COUNT 23 a 3 c 6 g 22 t
 ORIGIN
 Query Match 1.3%; Score 23.8; DB 8; Length 54;
 Best Local Similarity 72.1%; Pred. No. 1.2e+06;
 Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 1588 AACAGTACAACCTCTATCATGTTTGTATGCTTTTGTATATA 1630
 Db 54 AAAAAATACAAAATTATCATTTTTCATATGCTTTTTCGATA 12
 RESULT 21
 AX435571
 LOCUS
 DEFINITION Sequence 3986 from Patent WO0229113.
 Query Match 1.3%; Score 23.8; DB 8; Length 54;
 Best Local Similarity 72.1%; Pred. No. 1.2e+06;
 Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 1588 AACAGTACAACCTCTATCATGTTTGTATGCTTTTGTATATA 1630
 Db 54 AAAAAATACAAAATTATCATTTTTCATATGCTTTTTCGATA 12

```

ACCESSION AX435571
VERSION AX435571.1 GI:21660379
SOURCE Bacillus licheniformis
ORGANISM Bacillus licheniformis
REFERENCE 1
AUTHORS Berka, R. and Clausen, I. G.
TITLE Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 3986 11-APR-2002;
Novozymes Biotech, Inc. (US); Novozymes A/S (DK)
FEATURES
    source
        1..87
            /organism="Bacillus licheniformis"
            /mol_type="genomic DNA"
            /db_xref="taxon:11402"
BASE COUNT 15 a 18 c 25 g 29 t
ORIGIN
Query Match 1.3%; Score 23.8; DB 6; Length 87;
Best Local Similarity 59.7%; Pred. No. 1.2e+06;
Matches 40; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 331 GAGCGTATATTAGATTGGAGCGTCATCGCTCGGAGCTGTCTCTGTAATCTTT 390
Db 1 GAGCGAGGTACGGACACGAGCGTTTGCTTAATTGTTCTGTTCTATCTCTT 60

QY 391 GCTTATC 397
Db 61 AATCATC 67

RESULT 22
AF187253
LOCUS
DEFINITION Polytrichadelphus magellanicus trnL-trnF intergenic spacer,
chloroplast sequence.
ACCESSION AF187253
VERSION AF187253.1 GI:15592942
KEYWORDS chloroplast Polytrichadelphus magellanicus
ORGANISM Polytrichadelphus magellanicus
REFERENCE 1
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Bryophyta;
Polytrichopsida; Polytrichales; Polytrichaceae; Polytrichadelphus.
TITLE 1 (bases 1 to 87)
JOURNAL Molecular divergence of New Zealand samples of Meteoriaceae and
REFERENCE 2 (bases 1 to 87)
AUTHORS Lembohyllaceae
JOURNAL Unpublished
AUTHORS Quandt, D.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-1999) Institut fuer Biologie-Systematische
Botanik und Pflanzensoziographie, Freie Universitaet Berlin,
Altensteinsrasse 6, Berlin 14195, Germany
FEATURES
    source
        1..87
            /organism="Polytrichadelphus magellanicus"
            /organelle="plastid:chloroplast"
            /mol_type="genomic DNA"
            /db_xref="taxon:111667"
    misc_feature 1..87
        /note="trnL-trnF intergenic spacer"
BASE COUNT 42 a 3 c 6 g 36 t
ORIGIN
Query Match 1.3%; Score 23.8; DB 8; Length 87;
Best Local Similarity 62.7%; Pred. No. 1.2e+06;
Matches 37; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1033 AACATTGGATCTTCAGAAAGATCAAAATCTATTTTAGTGATPAAAGCTTTAGCTATA 1091
Db 1 AAAATTATATATAAGAAAAATAAATTTTCATTTTATTGACATAAGTTAATTTTAA 59

RESULT 23
ATH526349/c
LOCUS
DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
118E10.
ACCESSION AJ526349
VERSION AJ526349.1 GI:26794609
SOURCE left border; T-DNA flanking sequence.
ORGANISM Arabidopsis thaliana (thale cress)
REFERENCE 1
AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,
Lepiniec, L., Caboche, M. and Lecharny, A.
TITLE T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE 22363535
PUBMED 12446565
REFERENCE 2 (bases 1 to 93)
AUTHORS Balzergue, S.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
FEATURES
    source
        1..93
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /cultivar="Wassiljewskaja"
            /db_xref="taxon:3702"
            /clone="118E10"
            /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
    misc_feature 1..93
        /note="T-DNA flanking sequence"
        left border"
BASE COUNT 47 a 8 c 11 g 27 t
ORIGIN
Query Match 1.3%; Score 23.8; DB 8; Length 93;
Best Local Similarity 53.8%; Pred. No. 1.2e+06;
Matches 49; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 816 AAAAATTGGCTCTTTCTGGAAGTGTCTTTCTTTAGTTGCGTCTCTCGGGTTG 875
Db 93 AAAACAATCTCTGCTTTTCACTATTTTCTTTCTTTTATTAGTTTACGCCGTT 34

QY 876 GAGCTGGAATTCGTTTGGGAATGGATT 906
Db 33 GTTATTGAATTGATATTAAATAAATATT 3

RESULT 24
LACPIAA/c
LOCUS
DEFINITION L.lactis P1 promoter, partial sequence.
ACCESSION M63184
VERSION M63184.1 GI:149458

```

KEYWORDS
SOURCE Lactococcus lactis
ORGANISM Lactococcus lactis
 Bacteria: Firmicutes; Lactobacillales; Streptococcaceae;
 Lactococcus.
 1 (bases 1 to 100)
REFERENCE Koivula, T., Sibakov, M. and Palva, I.
AUTHORS Isolation and Characterization of Lactococcus lactis subsp. lactis
TITLE promoters
JOURNAL Appl. Environ. Microbiol. 57 (2), 333-340 (1991)
MEDLINE 91197100
PUBMED 1707605
COMMENT Original source text: L.lactis, cDNA to mRNA.
FEATURES Location/Qualifiers
 source
 1..100
 /organism="Lactococcus lactis"
 /mol_type="genomic DNA"
 /db_xref="taxon:1358"
 1..100
 /partial
 /note="pl"
 BASE COUNT 43 a 11 c 17 g 29 t
 ORIGIN
 Query Match 1..3%; Score 23.8; DB 1; Length 100;
 Best Local Similarity 59.7%; Pred. No. 1.2e+06;
 Matches 40; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 QY 1389 TAGCACATTTCTTTTGAGAAAATTCAAATAGATTGCAAGCCATTGGTCTTAATGA 1448
 DB 94 TCGGACTTATGATTATTTAGCATTTTTCACAAATGCAAGTCTTTCTTCATAAAGC 35
 QY 1449 CTGTATT 1455
 DB 34 CTTTATT 28
 RESULT 25
 AX485729 65 bp DNA linear PAT 16-AUG-2002
 LOCUS Sequence 3029 from Patent WO02053728.
DEFINITION
ACCESSION AX485729
VERSION AX485729.1 GI:22319945
KEYWORDS
SOURCE Candida albicans
ORGANISM Candida albicans
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K.L.
AUTHORS Gene disruption methodologies for drug target discovery
TITLE Patent: WO 02053728-A 3029 11-JUL-2002;
JOURNAL Eliira Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
 source
 1..65
 /organism="Candida albicans"
 /mol_type="genomic DNA"
 /db_xref="taxon:5476"
 BASE COUNT 21 a 7 c 18 g 19 t
 ORIGIN
 Query Match 1..3%; Score 23.6; DB 6; Length 65;
 Best Local Similarity 69.6%; Pred. No. 1.3e+06;
 Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 268 AAAACAATGTACAAAGGCAATTCAGCATGCTGTGGCTGGCA 313
 DB 19 AAAGCAATGTATAAAGGCTTGTATAGGATGCTGTACATTAGCA 64
 RESULT 26
 MIAATVR/c 72 bp DNA linear INV 11-JUN-2003
 LOCUS MIAATVRV

DEFINITION Mosquito mitochondrial tRNA-Val (UAC).
ACCESSION X03925
VERSION X03925.1 GI:12641
KEYWORDS transfer RNA; transfer RNA-Val.
SOURCE mitochondrion Aedes albopictus (Asian tiger mosquito)
ORGANISM Aedes albopictus
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
REFERENCE 1 (bases 1 to 72)
AUTHORS Dubin, D.T., HsuChen, C.C. and Tillotson, L.E.
TITLE Mosquito mitochondrial transfer RNAs for valine, glycine and
 glutamate: RNA and gene sequences and vicinal genome organization
JOURNAL Curr. Genet. 10 (9), 701-707 (1986)
MEDLINE 88194654
COMMENT Data kindly reviewed (06-OCT-1986) by D. Dubin.
FEATURES Location/Qualifiers
 source
 1..72
 /organism="Aedes albopictus"
 /organelle="mitochondrion"
 /mol_type="genomic DNA"
 /db_xref="taxon:7160"
 1..72
 /product="tRNA-Val"
 modified_base 9
 /note="1-methyladenosine"
 /mod_base=m1a
 modified_base 10
 /note="2-methylguanosine"
 /mod_base=m2g
 modified_base 27
 /note="pseudouridine"
 /mod_base=p
 modified_base 28
 /note="pseudouridine"
 /mod_base=p
 misc_feature 34..36
 /note="anticodon UAC"
 modified_base 37
 /note="N-(9-beta-D-ribofuranosylpurine-6-yl) carbamoyl)
 threonine in tRNA-Val"
 modified_base 38
 /mod_base=OTHER
 modified_base 39
 /note="pseudouridine"
 /mod_base=p
 modified_base 64
 /note="pseudouridine"
 /mod_base=p
 modified_base 65
 /note="pseudouridine"
 /mod_base=p
 BASE COUNT 31 a 8 g 27 t
 ORIGIN
 Query Match 1..3%; Score 23.6; DB 3; Length 72;
 Best Local Similarity 61.3%; Pred. No. 1.3e+06;
 Matches 38; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
 QY 662 TTGACATTGATTTAGGAATTTCTCTCAATGTTTACGGCTCTTTTCATGACTAAATTT 721
 DB 68 TTTAAATGATTTGCACAAATTTCTTTCAATGTAATGAAATACTTTTACTAATTAAGCT 9
 QY 722 TT 723
 DB 8 TT 7
 RESULT 27
 BD035147 79 bp DNA linear PAT 27-AUG-2002
 LOCUS BD035147
DEFINITION Sequence tag and encoded human protein.

ACCESSION BD035147
VERSION BD035147.1 GI:22576889
KEYWORDS JP 2001269182-A/11393.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 79)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 11393 02-OCT-2001;
GENSET OS Homo sapiens (human)
COMMENT PN JP 2001269182-A/11393
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PI JEAN RAPUTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
CC G06F15/40
PH Key Location/Qualifiers.
source 1..79
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606" 32 t
BASE COUNT 22 a 11 c 14 g
ORIGIN
Query Match 1.3%; Score 23.6; DB 6; Length 79;
Best Local Similarity 69.6%; Pred. No. 1.3e+06;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1348 TTTAATTCATGACCTTTGGCTACCTGTCAGCTGTGTTATAGCA 1393
|||||
Db 3 TTTAATTCATGACCTTTGGCTACCTGTCAGCTGTGTTATAGCA 48
|||||
RESULT 28
AX241109
LOCUS AX241109 90 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 347 from Patent WO0160975.
ACCESSION AX241109
VERSION AX241109.1 GI:15797984
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Roemer,T., Jiang,B., Boone,C. and Bussey,H.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 0160975-A 347 23-AUG-2001;
Elitra Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1..90
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="DNA primer"
BASE COUNT 27 a 10 c 25 g 28 t
ORIGIN
Query Match 1.3%; Score 23.6; DB 6; Length 90;
Best Local Similarity 69.6%; Pred. No. 1.3e+06;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 268 AAACAATGTACACAGGCAATATCTCAGCATGCTGTGGCTGGCA 313
|||||
Db 19 AAACAATGTATAAAGGCTTGGTATAAGGATGCTGTACATTAGCA 64
|||||

RESULT 29
AX326604/c
LOCUS AX326604 98 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 25 from Patent WO0192579.
ACCESSION AX326604
VERSION AX326604.1 GI:18097367
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wenz,H.M. and Schroth,G.P.
TITLE Methods for detecting target nucleic acids using coupled ligation
JOURNAL Patent: WO 0192579-A 25 06-DEC-2001;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1..98
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="allele of SEQ ID NO: 24; pivotal nucleotide (49)"
BASE COUNT 23 a 15 c 20 g 40 t
ORIGIN
Query Match 1.3%; Score 23.6; DB 6; Length 98;
Best Local Similarity 58.6%; Pred. No. 1.3e+06;
Matches 41; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1156 TGTCAGAAACAGCGCTAGATTCTCTACGGAACCTCTAAACGAAACGAAATTTGGTC 1215
|||||
Db 77 TGTCAGAAACAGCGCTAGATTCTCTACGGAACCTCTAAACGAAACGAAATTTGGTC 18
|||||
QY 1216 AAAGGTAAGC 1225
|||||
Db 17 AAAGGTAAGC 8
|||||
RESULT 30
AX326665
LOCUS AX326665 98 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 86 from Patent WO0192579.
ACCESSION AX326665
VERSION AX326665.1 GI:18097428
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wenz,H.M. and Schroth,G.P.
TITLE Methods for detecting target nucleic acids using coupled ligation
JOURNAL Patent: WO 0192579-A 86 06-DEC-2001;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1..98
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="complement of SEQ ID NO: 25"
BASE COUNT 40 a 20 c 15 g 23 t
ORIGIN
Query Match 1.3%; Score 23.6; DB 6; Length 98;
Best Local Similarity 58.6%; Pred. No. 1.3e+06;
Matches 41; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1156 TGTCAGAAACAGCGCTAGATTCTCTACGGAACCTCTAAACGAAACGAAATTTGGTC 1215
|||||

Db 22 TGTACGACACCCATGTATCCATATACGGCAATCAAAATAGGAACCTATGACCTAAAGC 81

QY 1216 AAAGGTAAGC 1225

Db 82 AAAGGTAAC 91

RESULT 31
MMGGH3/c
LOCUS Mus musculus gamma-glutamyl hydrolase gene, exon 3.
DEFINITION AF090727.1 GI:5712212
ACCESSION AF090727.1
VERSION AF090727.1
KEYWORDS 3 of 8
SEGMENT Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 88)
AUTHORS Esaki, T., Roy, K., Yao, R., Galivan, J. and Sirotnak, F.M.
TITLE Cloning of mouse gamma-glutamyl hydrolase in the form of two cDNA variants with different 5' ends and encoding alternate leader peptide sequences
JOURNAL Gene 219 (1-2), 37-44 (1998)
MEDLINE 98434456
PUBMED 9756990
REFERENCE 2 (bases 1 to 88)
AUTHORS Esaki, T., Masumoto, N., Hayes, P., Chen, J. and Sirotnak, F.M.
TITLE Organization and structure of the mouse gamma-glutamyl hydrolase gene and the functional identification of its promoter
JOURNAL Gene 234 (1), 93-100 (1999)
MEDLINE 99321803
PUBMED 10393243
REFERENCE 3 (bases 1 to 88)
AUTHORS Masumoto, N., Esaki, T. and Sirotnak, F.M.
TITLE Additional organizational features of the murine gamma-glutamyl hydrolase gene. Two remotely situated exons within the complement C3 gene locus encode an alternate 5' end and proximal ORF under the control of a bidirectional promoter
JOURNAL Gene 268 (1-2), 183-194 (2001)
MEDLINE 21261955
PUBMED 11368914
REFERENCE 4 (bases 1 to 88)
AUTHORS Esaki, T. and Sirotnak, F.M.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1998) Molecular Therapeutics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA
FEATURES
source Location/Qualifiers
1. .88
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/SVJ"
/db_xref="taxon:10090"
20. .70
/number=3
21 a 11 c 21 g 35 t
BASE COUNT 21 a 11 c 21 g 35 t
ORIGIN
Query Match 1.3%; Score 23.4; DB 10; Length 88;
Best Local Similarity 63.2%; Pred. No. 1.4e+06;
Matches 36; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
714 CTAATTTTCTTCATCGTGTGATGAATAAGACCAACATACACAGTTGCATATGA 770
Db 58 CTAATAAGTTCTTACTCTCATCTAGTACAGTACAGGCTGAAACACAGGAATGA 2

left end.
J01518
J01518.1 GI:343852
AT-rich region; GC rich region.
1 of 2
Mitochondrion Saccharomyces cerevisiae (baker's yeast)
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 91)
AUTHORS de Zamaroczy, M., Faugeron-Ponty, G. and Bernardi, G.
TITLE Excision sequences in the mitochondrial genome of yeast
JOURNAL Gene 21 (3), 193-202 (1983)
MEDLINE 83210931
PUBMED 6343188
COMMENT Original source text: Yeast (Saccharomyces cerevisiae)
Mitochondrial DNA.
Additional sequences reported in [1], but sequenced in earlier papers, appear in separate entries. Excision repeat corresponds to bases 35 to 39.
FEATURES
source Location/Qualifiers
1. .91
/organism="Saccharomyces cerevisiae"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
52 a 0 c 2 g 37 t
BASE COUNT 52 a 0 c 2 g 37 t
ORIGIN
Query Match 1.3%; Score 23.4; DB 8; Length 91;
Best Local Similarity 55.6%; Pred. No. 1.4e+06;
Matches 45; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1009 TTCTCTAGAGACTTCGCTATTCAACATTTGGATCTTCAGAAAGATCAAAATCTATT 1068
Db 81 TTTATTTATATATTTATTTATTAATAATATTAATCTATTATTAATAATATATAT 22
QY 1069 TAGTGATAAAGCTTTAAAGCTA 1089
Db 21 TTAATATATTTCTTTAAATTA 1
RESULT 33
BD037934
LOCUS BD037934
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD037934
VERSION BD037934.1 GI:22579676
KEYWORDS JP 2001269182-A/14180.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 94)
AUTHORS Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 14180 02-OCT-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/14180
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FEATURES
source Location/Qualifiers
1. .94

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT      19 a      8 c      13 g      52 t      2 others
ORIGIN
Query Match      1.3%; Score 23.4; DB 6; Length 94;
Best Local Similarity 63.2%; Pred. No. 1.4e+06;
Matches 36; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1348 TTTAATTCATGACCTTTGGCTACCTGTGCGAGCTTTGTTTATAGCACATTTCTTTT 1404
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26 TTTTATGTTATCCTTTATGCACTTATAATGTTCTTTTATAGGATTTCTTTT 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 34
AX036605
LOCUS      Sequence 11 from Patent FR2791358.
ACCESSION AX036605
VERSION   AX036605.1 GI:11226200
KEYWORDS  synthetic construct
SOURCE    synthetic construct
ORGANISM  artificial sequences.
REFERENCE 1
AUTHORS   Rance, I., Theisen, M. and Gruber, V.
JOURNAL   Patent: FR 2791358-A 11 29-SEP-2000;
MERISTEM THERAPEUTICS (FR)
FEATURES
source    /organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="Directional building block SI"
BASE COUNT      26 a      18 c      10 g      16 t
ORIGIN
Query Match      1.2%; Score 23.2; DB 6; Length 70;
Best Local Similarity 61.7%; Pred. No. 1.6e+06;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1029 TTCAAACATTTGGATCTTCAGAAAGATCAAAATCTATTTCAGTATAAGCTTTAAGCT 1088
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 TTCAAACATACAAATTCAGTAGAGAACTATTACTCTTGAGAAACCTAGAGGAT 55
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 35
BD039479/c
LOCUS      Sequence tag and encoded human protein.
DEFINITION BD039479
ACCESSION BD039479.1 GI:22581221
VERSION   BD039479.1
KEYWORDS  JP 2001269182-A/15725.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 75)
AUTHORS   Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
TITLE     Sequence tag and encoded human protein
JOURNAL   Patent: JP 2001269182-A 15725 02-OCT-2001;
GENSET
COMMENT    OS Homo sapiens (human)
FN       JP 2001269182-A/15725
PD       02-OCT-2001
PF       24-FEB-2000 JP 2000118773
PR       26-FEB-1999 US 60/122487
PI       JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PC       C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC       C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC

```

```

G06F15/40
CC        Location/Qualifiers.
FH        Location/Qualifiers
FEATURES   source
            1..75
BASE COUNT      19 a      7 c      13 g      36 t
ORIGIN
Query Match      1.2%; Score 23.2; DB 6; Length 75;
Best Local Similarity 61.7%; Pred. No. 1.6e+06;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1160 AGAAACAGCGCTAGATTTCTTACGGAACCTTAAACGAGCAAAATTTGGTCAAAG 1219
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 AAAAAAATACATGATTTCTCTATGGATATGCCACATACCCATATATCTTAAAGTCAAAG 11
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 36
ATH530983
LOCUS      Arabidopsis thaliana T-DNA flanking sequence, left border, clone
DEFINITION 222F11.
ACCESSION AJ530983
VERSION   AJ530983.1 GI:26799243
KEYWORDS  left border; T-DNA flanking sequence.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS   Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,
Lepiniec, L., Caboche, M. and Lecharny, A.
TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE   22363535.
PUBMED    12446565
REFERENCE 2 (bases 1 to 86)
AUTHORS   Balzerque, S.
TITLE     Direct Submission
JOURNAL
COMMENT    Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'genoplante' (http://www.genoplante.com and
http://genoplante-info.inbio.gen.fr).
FEATURES   Location/Qualifiers
            1..86
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /culivar="Wassillewskija"
            /db_xref="taxon:3702"
            /clone="222F11"
            /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
            /note="T-DNA flanking sequence"
            left border"
BASE COUNT      46 a      7 c      9 g      24 t
ORIGIN
Query Match      1.2%; Score 23.2; DB 8; Length 86;

```


BD115800/c	BD115800	98 bp	DNA	linear	PAT 18-SEP-2002							
LOCUS	EST and encoded human protein.											
DEFINITION	BD115800											
ACCESSION	BD115800.1	GI:23210704										
VERSION	JP 2002010789-A/7877.											
KEYWORDS	Homo sapiens (human)											
SOURCE	Homo sapiens											
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											
REFERENCE	1 (bases 1 to 98)											
AUTHORS	Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.											
TITLE	EST and encoded human protein											
JOURNAL	Patent: JP 2002010789-A 7877 15-JAN-2002;											
	GENSET CORP											
COMMENT	OS Homo sapiens (human)											
	PN JP 2002010789-A/7877											
	PD 15-JAN-2002											
	PF 07-AUG-2000 JP 2000280989											
	PR 05-AUG-1999 US 60/147499											
	PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI											
	GIORDANO											
	PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC											
	C12N1/21,											
	PC C12N5/10, C12P21/02, C12P21/08, C12Q1/69, C12N15/00, C12N5/00, PC											
	C12N15/00											
	CC EST and encoded human protein											
	EH Key	Location/Qualifiers										
	FT source	1..98										
	FT	/Organism='Homo sapiens (human)'										
FEATURES	Location/Qualifiers											
source	1..98											
	/organism="Homo sapiens"											
	/mol_type="genomic DNA"											
	/db_xref="taxon:9606"											
BASE COUNT	82 a	3 c	6 g	7 t								
ORIGIN												
	Query Match 1.2%; Score 23.2; DB 6; Length 98;											
	Best Local Similarity 53.3%; Pred. No. 1.6e+06;											
	Matches 49; Conservative 0; Mismatches 43; Indels 0; Gaps 0;											
Qy	1608	TTTTGTTAAATGCTTTTGGTTTATAGCGGCTCCTCTCTCTTTAATTTGCATTTATTATGA	1667									
Db	95	TTT	36									
Qy	1668	CCATAGGGAATCTCTTAGAACCTTATCGTCT	1699									
Db	35	TTTTTCTCTCTTTTACAAAGACTTTATTGCTCT	4									
RESULT 44												
AY144376												
LOCUS	AY144376	99 bp	DNA	linear	VRL 07-MAY-2003							
DEFINITION	Emiliana huxleyi virus isolate OTU3 major capsid protein (MCP)											
TITLE	gene, partial cds.											
ACCESSION	AY144376											
VERSION	AY144376.1	GI:24429761										
KEYWORDS												
SOURCE	Emiliana huxleyi virus											
ORGANISM	Emiliana huxleyi virus											
REFERENCE	Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae.											
AUTHORS	1 (bases 1 to 99)											
TITLE	Schroeder,D.C., Oke,J., Hall,M., Malin,G. and Wilson,W.H.											
JOURNAL	Virus Succession Observed during an Emiliana huxleyi Bloom											
PUBLISHED	Appl. Environ. Microbiol. 69 (5), 2484-2490 (2003)											
REFERENCE	12732512											
AUTHORS												


```

source      1. .99
/organism="Emilia huxleyi virus"
/mol_type="genomic DNA"
/isolate="OTU3"
/db_xref="taxon:181208"
gene        complement(<1. .>99)
            /gene="MCP"
            complement(<1. .>99)
CDS         /gene="MCP"
            /codon_start=1
            /product="major capsid protein"
            /protein_id="AA52850.1"
            /db_xref="GI:24425762"
            /translation="NQPTTQVQGASHITVNRQGLLSMMYLKIV"
            28 a      23 c      23 g      25 t
BASE COUNT
ORIGIN

```

RESULT 45	AY144377	99 bp	DNA	linear	VRL 07-MAY-2003
LOCUS	AY144377	Emiliania huxleyi virus isolate OTU4 major capsid protein (MCP)			
DEFINITION	gene, partial cds.				
ACCESSION	AY144377.1	GI:24429763			
VERSION					
KEYWORDS					
SOURCE	Emiliania huxleyi virus				
ORGANISM	Emiliania huxleyi virus				
REFERENCE	Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae.				
AUTHORS	1 (bases 1 to 99)				
TITLE	Schroeder,D.C., Oke,J., Hall,M., Malin,G. and Wilson,W.H.				
JOURNAL	Virus Succession Observed during an Emiliania huxleyi Bloom				
PUBLISHED	Appl. Environ. Microbiol. 69 (5), 2484-2490 (2003)				
REFERENCE	12732512				
AUTHORS	2 (bases 1 to 99)				
TITLE	Schroeder,D.C., Oke,J., Hall,M., Evans,C., Malin,G. and Wilson,W.H.				
JOURNAL	Direct Submissions				
FEATURES	Submitted (26-AUG-2002) Marine Biological Association, Citadel Hill, Plymouth, Devon PL1 2PB, England				
source	Location/Qualifiers				
	1..99				
	/organism="Emiliania huxleyi virus"				
	/mol_type="genomic DNA"				
	/isolate="OTU4"				
gene	/db_xref="taxon:181208"				
	complement(<1..>99)				
	/gene="MCP"				
CDS	complement(<1..>99)				
	/gene="MCP"				
	/codon_start=1				
	/product="major capsid protein"				
	/protein_id="AA52851.1"				
	/db_xref="GI:24429764"				
	/translation="NQPTTQVQFGAESHTVNRQGDLLSNMYLKIV"				
BASE COUNT	26 a	23 c	25 g	24 t	1 others
ORIGIN					

```

Qy 59 GCRAACGTCGATCGCGTATGGCTGTAGTATTGACGGTTATATGGTCAGAGCCCTATT 118
    |||||
Db 13 GTACATCCAGCATAGCAATCGCCTTTCAGCGTTACGGTAATATGCGACTAGCCCCCAA 72
    |||||

Qy 119 TTAAACGT 126
    |||||
Db 73 CTGAACCT 80
    |||||

Search completed: October 23, 2003, 08:19:54
Job time : 6723 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 06:19:14 ; Search time 508 Seconds
(without alignments)
9840.309 Million cell updates/sec

Title: US-09-868-987-1

Perfect score: 1864

Sequence: 1 atggacttcgcgattgttc.....atgacaaattcagataatgc 1864

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 1287486

Minimum DB seq length: 5

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25.4	1.4	88	10	US-09-983-965-5388 Sequence 5388, Ap
C 2	24.2	1.3	65	12	US-09-908-975-23155 Sequence 23155, A
C 3	24	1.3	65	12	US-10-032-585-1037 Sequence 1037, Ap
C 4	23.8	1.3	87	10	US-09-974-300-3986 Sequence 3986, Ap
C 5	23.6	1.3	60	12	US-09-908-975-7106 Sequence 7106, Ap
C 6	23.6	1.3	65	12	US-10-032-585-3029 Sequence 3029, Ap
C 7	23.6	1.3	98	12	US-10-308-891-25 Sequence 25, Appl
C 8	23.6	1.3	98	12	US-10-308-891-86 Sequence 86, Appl
C 9	23.4	1.3	60	12	US-09-908-975-5572 Sequence 5572, Ap
C 10	23.2	1.2	60	12	US-09-908-975-18757 Sequence 18757, A
C 11	23.2	1.2	60	12	US-09-908-975-22233 Sequence 22233, A
C 12	23.2	1.2	70	11	US-09-963-803-34 Sequence 34, Appl
C 13	23.2	1.2	96	11	US-09-851-486-67 Sequence 67, Appl
C 14	23.2	1.2	96	12	US-10-223-666-156 Sequence 156, App
C 15	23	1.2	77	9	US-09-864-761-29497 Sequence 29497, A
C 16	22.8	1.2	65	12	US-10-032-585-560 Sequence 560, App

C 17	22.6	1.2	60	12	US-10-334-672-2	Sequence 2, Appli
C 18	22.6	1.2	60	12	US-09-908-975-21214	Sequence 21214, A
C 19	22.6	1.2	60	12	US-09-908-975-23155	Sequence 23155, A
C 20	22.6	1.2	60	12	US-10-334-671-2	Sequence 2, Appli
C 21	22.6	1.2	65	12	US-10-032-585-1037	Sequence 1037, Ap
C 22	22.6	1.2	91	13	US-10-092-140-5	Sequence 5, Appli
C 23	22.6	1.2	94	12	US-10-029-386-21155	Sequence 21155, A
C 24	22.6	1.2	96	11	US-09-851-486-61	Sequence 61, Appl
C 25	22.6	1.2	96	12	US-10-223-666-150	Sequence 150, App
C 26	22.6	1.2	100	11	US-09-849-928-15	Sequence 15, Appl
C 27	22.6	1.2	100	14	US-10-066-960-15	Sequence 15, Appl
C 28	22.4	1.2	65	12	US-09-908-975-27834	Sequence 27834, A
C 29	22.4	1.2	65	12	US-10-032-585-2537	Sequence 2537, Ap
C 30	22.4	1.2	96	9	US-09-864-761-18576	Sequence 18576, A
C 31	22.4	1.2	88	10	US-09-896-888A-15	Sequence 15, Appl
C 32	22.2	1.2	91	14	US-10-137-036-47	Sequence 47, Appl
C 33	22.2	1.2	93	10	US-09-960-352-4506	Sequence 4506, Ap
C 34	22	1.2	65	12	US-10-032-585-1310	Sequence 1310, Ap
C 35	22	1.2	96	10	US-09-878-574-7267	Sequence 7267, Ap
C 36	22	1.2	97	10	US-09-969-373-900	Sequence 900, App
C 37	22	1.2	98	9	US-09-864-761-21839	Sequence 21839, A
C 38	22	1.2	98	12	US-10-308-891-24	Sequence 24, Appl
C 39	21.8	1.2	65	12	US-09-908-975-1635	Sequence 1635, Ap
C 40	21.8	1.2	65	12	US-09-908-975-4780	Sequence 4780, Ap
C 41	21.8	1.2	65	12	US-10-032-585-586	Sequence 586, App
C 42	21.8	1.2	65	12	US-10-032-585-2664	Sequence 2664, Ap
C 43	21.8	1.2	87	14	US-10-104-818-13	Sequence 13, Appl
C 44	21.8	1.2	92	9	US-09-864-761-27613	Sequence 27613, A
C 45	21.8	1.2	92	14	US-10-005-338B-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-09-983-965-5388/c

; Sequence 5388, Application US/09983965

; Patent No. US20020137160A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

; FILE REFERENCE: 37-21(10297)C

; CURRENT APPLICATION NUMBER: US/09/983,965

; CURRENT FILING DATE: 2001-10-26

; PRIOR APPLICATION NUMBER: US 09/465,231

; PRIOR FILING DATE: 1999-12-15

; PRIOR APPLICATION NUMBER: US 60/113,678

; PRIOR FILING DATE: 1998-12-17

; NUMBER OF SEQ ID NOS: 5912

; SEQ ID NO 5388

; LENGTH: 88

; TYPE: DNA

; ORGANISM: Bos taurus

; FEATURE:

; OTHER INFORMATION: Clone ID: 43-LIB34-027-Q1-E1-C4

US-09-983-965-5388

Query Match 1.4%; Score 25.4; DB 10; Length 88;

Best Local Similarity 58.7%; Pred. No. 7.7e+03;

Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 211 GAGTCTTTTGTTCGAGGTTCTCAGTGAAGAGAGATCTCTTCTGATCTTGGGAAAAA 270

Db 88 GAGGCCACAGGTTATGAGATATCAGAGGAGAGAGAGATGTTGGTGGGTATA 29

QY 271 ACAATGTACACAGG 285

Db 28 TCAATGCAAAAAAGG 14

```
RESULT 2
US-09-908-975-29830
; Sequence 29830, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29830
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-29830

Query Match      1.3%; Score 24.2; DB 12; Length 65;
Best Local Similarity 66.0%; Pred. No. 1.4e+04;
Matches 35; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 816 AAAAATTGGCGTCTCTTTCTGGAAGTCTTTTCTTTAGTTCGGTGTCTC 868
Db 13 AAAGAAATTGGCCTGGATTGGTTGAGTTATCCCATTAAGCATCTTGTGTC 65

RESULT 3
US-10-032-585-3503/c
; Sequence 3503, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3503
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-3503

Query Match      1.3%; Score 24; DB 12; Length 65;
Best Local Similarity 60.9%; Pred. No. 1.5e+04;
Matches 39; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1684 AGGAATTTATCTCTCTTTATATTCACCACCTCTGTGTTGTTTATGGTCCGTAAGA 1743
Db 65 ATGGAATTTAGCCTCTCATCATATTTTAAACGCGTTGATGTTGATGAAGGATTATCAATA 6

QY 1744 AAAT 1747
Db 5 GACT 2

RESULT 4
US-09-974-300-3986
; Sequence 3986, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085 500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3986
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-3986

Query Match      1.3%; Score 23.8; DB 10; Length 87;
Best Local Similarity 59.7%; Pred. No. 2.1e+04;
Matches 40; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 331 GAGCGTATATTATAGATTGGAGGCGTCATCGCTTCGGGCGCTGTTCTTCTGAATCTTT 390
Db 1 GAGCGGAGGTACGGACACGGAAGCGGTTTTCCTTTAATGTTGTCTGTTCTATCTCTT 60

QY 391*GCTTATC 397
Db 61 AATCATC 67

RESULT 5
US-09-908-975-7106/c
; Sequence 7106, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7106
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-7106

Query Match      1.3%; Score 23.6; DB 12; Length 60;
Best Local Similarity 76.3%; Pred. No. 1.9e+04;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 447 CTGGATTGTTCTTGCTATGGGATGGCGGTAGATGCA 484
Db 60*CTTGACTGCTCTGGCTATGGAAAGGCGGTAGATCCA 23

RESULT 6
US-10-032-585-3029
```

; TITLE OF INVENTION: Methods for Detecting Target Nucleic Acids Using Coupled Ligation
 ; TITLE OF INVENTION: and Amplification
 ; FILE REFERENCE: 7414.0020-01000
 ; CURRENT APPLICATION NUMBER: US/10/308,891
 ; CURRENT FILING DATE: 2002-12-02
 ; PRIOR APPLICATION NUMBER: US 09/584,905
 ; PRIOR FILING DATE: 2000-05-30
 ; PRIOR APPLICATION NUMBER: US 09/724,755
 ; PRIOR FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 94
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 85
 ; LENGTH: 98
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: complement of SEQ ID NO: 25
 US-10-308-891-86

Query Match 1.3%; Score 23.6; DB 12; Length 98;
 Best Local Similarity 58.6%; Pred. No. 2.5e+04;
 Matches 41; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1156 TGTCAAGAACGCGCTAGATTCTCTACGGAACCTCTAAACGAAACGCAAAATTTTGGTC 1215
 DB 22 TGTCAACGACGCCATGTATCCATATACGGCAATCAAAATGAAGGAACCTTATGACCTAAGC 81
 QY 1216⁸ AAAGGTAAAGC 1225
 DB 82 AAAGGTAAAC 91

RESULT 9
 US-09-908-975-5572
 ; Sequence 5572, Application US/09908975
 ; Publication No. US20030165843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHOSHAN, Avi
 ; APPLICANT: WASSERMAN, Alon
 ; APPLICANT: MINTZ, Eli
 ; APPLICANT: MINTZ, Liat
 ; APPLICANT: FAIGLER, Simchon
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
 ; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
 ; FILE REFERENCE: 36688-0005
 ; CURRENT APPLICATION NUMBER: US/09/908,975
 ; CURRENT FILING DATE: 2001-07-20
 ; PRIOR APPLICATION NUMBER: US 60/287,724
 ; PRIOR FILING DATE: 2001-05-02
 ; PRIOR APPLICATION NUMBER: US 60/221,607
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 32337
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5572
 ; LENGTH: 60
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-908-975-5572

Query Match 1.3%; Score 23.4; DB 12; Length 60;
 Best Local Similarity 67.3%; Pred. No. 2.1e+04;
 Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 118 TTTAAACGTCCCATGAAATAATCATGCCAGTGTCTCAGGGAAATTTTACC 166
 DB 2³ TTTACCCCTCAAGTTTCAAGTTTCACTGTCTCAGAGAGGTTTTCC 50

RESULT 10
 US-09-908-975-18757
 ; Sequence 18757, Application US/09908975
 ; Publication No. US20030165843A1

```

; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT FILING DATE: 2001-07-20
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18757
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-908-975-18757

Query Match      1.2%; Score 23.2; DB 12; Length 60;
Best Local Similarity 70.5%; Pred. No. 2.4e+04;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 604 AGTATTGGCTCAGACTCTTTCTTCTTCCTAGATACAGGGCCTA 647
Db 11 AGGAATGGCTCAGCTCTTGTGCTGCTGCTTTAGGGGCCA 54

RESULT 11
US-09-908-975-22233/c
; Sequence 22233, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT FILING DATE: 2001-07-20
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22233
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-908-975-22233

Query Match      1.2%; Score 23.2; DB 12; Length 60;
Best Local Similarity 77.8%; Pred. No. 2.4e+04;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 995 CAGGAAGCTGGCTTCTTCTTCCTAGAGACTTCGGTATT 1030
Db 46 CAGGACGCTGCTCTTTGGGGCTTGGAGACTTCCTTACT 11

RESULT 12
US-09-963-803-34
; Sequence 34, Application US/09963803
; Publication No. US20030028922A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Meristem Therapeutics
; TITLE OF INVENTION: Chimeric expression promoters originating from commelina yellow mc
; FILE REFERENCE: 184332042
; CURRENT FILING DATE: 2001-09-26
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: PCT IB00/00370
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 70
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Directional Desoxynucleotide
; US-09-963-803-34

Query Match      1.2%; Score 23.2; DB 11; Length 70;
Best Local Similarity 61.7%; Pred. No. 2.7e+04;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1029 TTCAACATTTGGATCTTCAGAAAAGATCAAAATCTATTTAGTATAAAGCTTTAAGCT 1088
Db 6 TTCAACACATACAAATTCAGTAGAGAGAAACTCATTACTTTCGAAACCTAGAGGAT 65

RESULT 13
US-09-851-486-67
; Sequence 67, Application US/09851486
; Publication No. US20030036642A1
; GENERAL INFORMATION:
; APPLICANT: NEEOUJA JANJIC, LARRY GOLD
; TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR (PDGF) NUCLEIC
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESS: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,486
; FILING DATE: 08-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/991,743
; FILING DATE: 16-DECEMBER-1997
; APPLICATION NUMBER: 08/618,693
; FILING DATE: 20-MARCH-1996
; APPLICATION NUMBER: 08/479,783
; FILING DATE: 7-JUNE-1995
; APPLICATION NUMBER: 08/479,725
; FILING DATE: 7-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX66
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
```

; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/006661


```

; LENGTH: 65
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-1037

Query Match      1.2%; Score 22.6; DB 12; Length 65;
Best Local Similarity 75.1%; Pred. No. 3.7e+04;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY    295 AGCATGCTGGCTTGCCAAATGCTTATGTGTTTGAAG 331
       ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     63 AGAATAATGTTTGCCAAATGCTTTTGTAGAATG 27

RESULT 22
US-10-092-140-5/c
; Sequence 5, Application US/10092140
; Publication No. US20020164801A1
; GENERAL INFORMATION:
; APPLICANT: McGill University et al.
; TITLE OF INVENTION: HUMAN AND MAMMALIAN DNA
; REPLICATION ORIGIN CONSENSUS SEQUENCES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SWABEY OGILVY RENAULT
; STREET: 1981 McGill College Avenue - Suite 1600
; CITY: Montr,al
; STATE: QC
; COUNTRY: Canada
; ZIP: H3A 2Y3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/092,140
; FILING DATE: 06-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,750
; FILING DATE: 09-Jun-1999
; APPLICATION NUMBER: 60/033,374
; FILING DATE: 16-DEC-1996
; APPLICATION NUMBER: 60/047,322
; FILING DATE: 21-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: C't., France
; REGISTRATION NUMBER: 4166
; REFERENCE/DOCKET NUMBER: 1770-162PCT FC/1d
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 514 845-7126
; TELEFAX: 514 288-8389
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-092-140-5

Query Match      1.2%; Score 22.6; DB 13; Length 91;
Best Local Similarity 13.6%; Pred. No. 4.6e+04;
Matches 11; Conservative 47; Mismatches 23; Indels 0; Gaps 0;

QY    1591 AGTCACACTCTATCAGTTTTGTTTAATGCTTTTGTATTATAGCGGCCTCTCTCTTTAA 1650
       | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     90 AMYTKWMKMMWMMMTATSWTWMMKWCMRRYYTWAWCMMWAMTMMHKGTSTYDW 31
Ov    1651 TTTTTCGCAATTATTATGACCATT 1671

```



```

1 FILING DATE: 04-May-2001
2 CLASSIFICATION: <Unknown>
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: 08/952,793
5 FILING DATE: <Unknown>
6 APPLICATION NUMBER: 08/479,724
7 FILING DATE: 07-JUNE-1995
8 APPLICATION NUMBER: 08/472,256
9 FILING DATE: 07-JUNE-1995
10 APPLICATION NUMBER: 08/472,255
11 FILING DATE: 07-JUNE-1995
12 APPLICATION NUMBER: 08/477,829
13 FILING DATE: 07-JUNE-1995
14 ATTORNEY/AGENT INFORMATION:
15 NAME: Barry J. Swanson
16 REGISTRATION NUMBER: 33,215
17 REFERENCE/DOCKET NUMBER: NEX40C/PCT
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: (303) 793-3333
20 TELEFAX: (303) 793-3433
21 INFORMATION FOR SEQ ID NO: 15:
22 SEQUENCE CHARACTERISTICS:
23     LENGTH: 100 base pairs
24     TYPE: nucleic acid
25     STRANDEDNESS: single
26     TOPOLOGY: linear
27     MOLECULE TYPE: RNA
28     FEATURE:
29     OTHER INFORMATION: All C's are 2'-NH2 cytosine
30     FEATURE:
31     OTHER INFORMATION: All U's are 2'-NH2 uracil
32     SEQUENCE DESCRIPTION: SEQ ID NO: 15:
33 US-09-849-928-15
34
35 Query Match      1.2%; Score 22.6; DB 11; Length 100;
36 Best Local Similarity 68.9%; Pred.No. 4.8e+04;
37 Matches 31; Conservative 0; Mismatches 14; Indels
38
39 QY 1345 CGCTTTAATTCATGACCTTTTGGCTACCTGTGCAGCTCTGTAT 1389
40      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41 Db 59 CGTCTAATTCAGTACATTTCGATTCGCCGCCGAGTCTGTGTAT 15
42
43 RESULT 27
44 US-10-066-960-15/c
45 ; Sequence 15, Application US/10066960
46 ; Publication No. US20030049644A1
47 ; GENERAL INFORMATION:
48 APPLICANT: PARMA, et al.
49 TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
50 TO LECTINS
51 NUMBER OF SEQUENCES: 390
52 CORRESPONDENCE ADDRESS:
53 ADDRESSEE: Swanson & Bratschun, L.L.C.
54 STREET: 8400 E. Prentice Avenue, Suite 200
55 CITY: Englewood
56 STATE: Colorado
57 COUNTRY: USA
58 ZIP: 80111
59 COMPUTER READABLE FORM:
60 MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
61 * COMPUTER: IBM pc compatible
62 OPERATING SYSTEM: MS-DOS
63 SOFTWARE: WordPerfect 6.0
64 CURRENT APPLICATION DATA:
65 APPLICATION NUMBER: US/10/066,960
66 FILING DATE: 04-Feb-2002
67 CLASSIFICATION: <Unknown>
68 PRIOR APPLICATION DATA:
69 APPLICATION NUMBER: 08/952,793
70 FILING DATE: 1999-DEC-03
71 APPLICATION NUMBER: PCT/US96/09455
72 FILING DATE: 05-JUNE-1995

```

```

; APPLICATION NUMBER: 08/479,724
; FILING DATE: 07-JUNE-1995
; APPLICATION NUMBER: 08/472,256
; FILING DATE: 07-JUNE-1995
; APPLICATION NUMBER: 08/472,255
; FILING DATE: 07-JUNE-1995
; APPLICATION NUMBER: 08/477,829
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX40C/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All C's are 2'-NH2 cytosine
; FEATURE:
; OTHER INFORMATION: All U's are 2'-NH2 uracil
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-066-960-15
Query Match 1.2%; Score 22.6; DB 14; Length 100;
Best Local Similarity 68.9%; Pred. No. 4.8e+04;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1345 CGTTTAAATTCAGACCTTTGGCTACCTGTGACGCTGCTTTAT 1389
DB 59 CGTTCTAAATTCAGTACACTTTCGATTCCCGCGAGTCTGTGTAT 15

RESULT 28
US-09-908-975-27834
; Sequence 27834, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27834
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-27834
Query Match 1.2%; Score 22.4; DB 12; Length 65;
Best Local Similarity 66.7%; Pred. No. 4.2e+04;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 766 TATCATGAATAAGTTCGTGGGGTAAGCATGATTTCTTCAGAGGATG 813
DB 18 TAGGAACAATCAGTGATCCGGTGACAGCCTGGGTCTTTTAGAGGCTG 65

```

```

RESULT 29
US-10-032-585-2537/c
; Sequence 2537, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2537
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-2537
Query Match 1.2%; Score 22.4; DB 12; Length 65;
Best Local Similarity 62.5%; Pred. No. 4.2e+04;
Matches 35; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 680 ATTTCTCTCAATGTTTACGGCTCTTTTCATGACTAAATTTTCTTCATGCTGTG 735
DB 57 ATTCITTTTGAATACTTTCGTTTCGTCGATCTTTTTCGCTTAATAGTGTG 2

RESULT 30
US-09-864-761-18576
; Sequence 18576, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```

PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 18576
 LENGTH: 96
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC009484.1
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.1
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.4
 OTHER INFORMATION: EST HUMAN HIT: AL038099.2, EVALUE 2.00e-26
 OTHER INFORMATION: SWISSPROT HIT: P11369, EVALUE 2.00e-04
 OTHER INFORMATION: NT HIT: A0229043.1, EVALUE 4.00e-21
 US-09-864-761-18576

 Query Match 1.2%; Score 22.4; DB 9; Length 96;
 Best Local Similarity 62.5%; Pred. No. 5.3e+04;
 Matches 35; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

 QY 688 TTCAATGTTTACGGCTTTTCATGACTAAATTTTCTTCATGCTCTGGATGAATA 743
 DB 31 TGGCATGATTTTGTCTTTTATGCTGAGTAGTATCCATGGTGTGTGTATA 86

 RESULT 31
 US-09-896-888A-15/c
 Sequence 15, Application US/09896888A
 Patent No. US20020116723A1
 GENERAL INFORMATION:
 APPLICANT: The University of British Columbia
 TITLE OF INVENTION: Insect Expression Vectors
 FILE REFERENCE: 80021-44
 CURRENT APPLICATION NUMBER: US/09/896,888A
 CURRENT FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US/09/048,911
 PRIOR FILING DATE: 1998-03-26
 PRIOR APPLICATION NUMBER: 60/049,946
 PRIOR FILING DATE: 1997-03-27
 NUMBER OF SEQ ID NOS: 50
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 15
 LENGTH: 88
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Fragment of
 OTHER INFORMATION: promoter sequence of the AcMNPV ien gene
 US-09-896-888A-15

 Query Match 1.2%; Score 22.2; DB 10; Length 88;
 Best Local Similarity 54.2%; Pred. No. 5.7e+04;
 Matches 45; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

 QY 439 AGGACTCGTGGATTTCTTCTCTATGGGATGGCGGTAGATGCAAAATTTCTTGATT 498
 DB 86 ACCGATTGTAGAGATTGTACTGTATATGGAGTGTCTAGGCAAAAGTGAACCTTTTTCATT 27

QY 499 CGAAAGAAATCCGAGAGGAATTTT 521
 DB 26 GCAGAAAAAATTCATTTTAAATTT 4

 RESULT 32
 US-10-137-036-47/c
 Sequence 47, Application US/10137036
 Publication No. US20030101478A1
 GENERAL INFORMATION:
 APPLICANT: Perera, Ranjan
 APPLICANT: Rice, Stephen
 APPLICANT: Eagleton, Claire
 APPLICANT: Lasham, Annette
 APPLICANT: Wood, Marion
 APPLICANT: Visser, Elizabeth
 TITLE OF INVENTION: Compositions and Methods for the
 TITLE OF INVENTION: Modification of Gene Expression
 FILE REFERENCE: 11000.1036c4
 CURRENT APPLICATION NUMBER: US/10/137,036
 CURRENT FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: PCT/NZ 01/00115
 PRIOR FILING DATE: 2001-06-20
 PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/724,624
 PRIOR FILING DATE: 2000-11-28
 PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/598,401
 PRIOR FILING DATE: 2000-06-20
 PRIOR APPLICATION NUMBER: PCT/NZ00/00018
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 60/146,591
 PRIOR FILING DATE: 1999-07-30
 PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/276,599
 PRIOR FILING DATE: 1999-03-25
 NUMBER OF SEQ ID NOS: 143
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 47
 LENGTH: 91
 TYPE: DNA
 ORGANISM: Pinus radiata
 US-10-137-036-47

 Query Match 1.2%; Score 22.2; DB 14; Length 91;
 Best Local Similarity 64.7%; Pred. No. 5.9e+04;
 Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

 QY 797 GATTTCTTCAGAGGATGCAAAAACTTTGGGCTGTTCTCGGAAGTGTTTT 847
 DB 75 GAATTTCTTCGGAGAACCAAGACTGTGGCCATTTACCGGCTTTGTAGTT 25

 RESULT 33
 US-09-960-352-4506
 Sequence 4506, Application US/09960352
 Patent No. US20020137139A1
 GENERAL INFORMATION:
 APPLICANT: Warren, Wesley C.
 APPLICANT: Tao, Nengbing
 APPLICANT: Byatt, John C.
 APPLICANT: Mathialagan, Nagappan
 TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
 FILE REFERENCE: 16511.006/37-21(10298)C
 CURRENT APPLICATION NUMBER: US/09/960,352
 CURRENT FILING DATE: 2001-09-24
 NUMBER OF SEQ ID NOS: 15112
 SEQ ID NO 4506
 LENGTH: 93
 TYPE: DNA
 ORGANISM: Bos taurus
 OTHER INFORMATION: Clone ID: 20-BOWMS1-022-Q1-E1-E7
 US-09-960-352-4506

 Query Match 1.2%; Score 22.2; DB 10; Length 93;

Qy	1071	GT	1072
Db	66	GT	67

RESULT 37

US-09-864-761-21839

Sequence 21839, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aescmica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,365

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

QY 217 TTTTGTCCGAGTCTCAGTGAAGAGCATCTCTTCTGATCTGGG 265
 Db 60 TTTCTGCTCTTTGGTCCCTGTGAAGCTACTCTGTGTCTGTGATGAG 12

RESULT 41
 US-10-032-585-586/c
 ; Sequence 586, Application US/10032585
 ; Publication No. US20030180953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Terry, Roemer D.
 ; APPLICANT: Bo, Jiang
 ; APPLICANT: Charles, Boone
 ; APPLICANT: Howard, Bussey
 ; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
 ; FILE REFERENCE: 10182-005-999
 ; CURRENT APPLICATION NUMBER: US/10/032,585
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 8000
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 586
 ; LENGTH: 65
 ; TYPE: DNA
 ; ORGANISM: Candida albicans
 US-10-032-585-586

Query Match 1.2%; Score 21.8; DB 12; Length 65;
 Best Local Similarity 61.4%; Pred. No. 6.1e+04;
 Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
 QY 620 CTTCTTTCTTCAGATACAGGCTATTAAAGGGTCTTGTGACATTGATTTA 676
 Db 65 CCTTATGTTTCAAGGATTTGGATATATAGGTTTGGATTGGAAATGATTTA 9

RESULT 42
 US-10-032-585-2664
 ; Sequence 2664, Application US/10032585
 ; Publication No. US20030180953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Terry, Roemer D.
 ; APPLICANT: Bo, Jiang
 ; APPLICANT: Charles, Boone
 ; APPLICANT: Howard, Bussey
 ; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
 ; FILE REFERENCE: 10182-005-999
 ; CURRENT APPLICATION NUMBER: US/10/032,585
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 8000
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2664
 ; LENGTH: 65
 ; TYPE: DNA
 ; ORGANISM: Candida albicans
 US-10-032-585-2664

Query Match 1.2%; Score 21.8; DB 12; Length 65;
 Best Local Similarity 61.4%; Pred. No. 6.1e+04;
 Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
 QY 488 GTTCTTGTTATCCAAAGATCCGAGAGGATTTTATTTGTTCTCAAGTCTTAAAAA 544
 Db 7 GTCTTTGTTTCCAACTTCTTGGAAAGATTTTTTCTGTGACAAAGTATTATCCAA 63

RESULT 43
 US-10-104-818-13/c
 ; Sequence 13, Application US/10104818
 ; Publication No. US20030087398A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vaxigenics, Inc.
 ; TITLE OF INVENTION: A Method for Analyzing Polynucleotides

; FILE REFERENCE: 265/034
 ; CURRENT APPLICATION NUMBER: US/10/104,818
 ; CURRENT FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: 09/394,774
 ; PRIOR FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 87
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hypothetical sequence to demonstrate application..
 US-10-104-818-13

Query Match 1.2%; Score 21.8; DB 14; Length 87;
 Best Local Similarity 65.3%; Pred. No. 7.3e+04;
 Matches 32; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 1334 AGTGGCGTATGCGCTTTAATTCATGACCTTTTGGCTACTGTGCAGTCT 1382
 Db 69 AGGACATATGAATTTTCATTCAGCAGCTTGATGTGCGGGAAGTCT 21

RESULT 44
 US-09-864-761-27613
 ; Sequence 27613, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aecmica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203

Search completed: October 23, 2003, 09:33:10
Job time : 508 secs

;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 27613
;; LENGTH: 92
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC004457.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.72
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.72
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.58
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.69
;; OTHER INFORMATION: EST HUMAN HIT: AA806461.1, EVALUE 1.80e-02
;; OTHER INFORMATION: NT HIT: AF045555.1, EVALUE 1.20e-02
US-09-864-761-27613

Query Match 1.2%; Score 21.8; DB 9; Length 92;
Best Local Similarity 56.2%; Pred. No. 7.6e+04;
Matches 41; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 584 TTGTGATTCTTAAGTACAGTATGGCTCAGCAGCTCTTTCTTCTAGATACAGGG 643
DB 1 TTGCTTTGACACTAACTAGTCTCTTGACATCAGCATCAGTTTCTCATCTAAAGGA 60
QY 644 CCTATTAAAGGT 656
DB 61 GTAGTTAATGTT 73

RESULT 45
US-10-005-338B-36/c
;; Sequence 36, Application US/10005338B
;; Publication No. US20030044895A1
;; GENERAL INFORMATION:
;; APPLICANT: DENEFFLE, Patrice
;; APPLICANT: ROSIER-MONTUS, Marie-Francoise
;; APPLICANT: PRADES, Catherine
;; APPLICANT: ARNOULD-REGUIGNE, Isabelle
;; APPLICANT: DUVERGER, Nicolas
;; APPLICANT: ALLIEMETS, Rando
;; APPLICANT: DEAN, Michael
;; TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENES
;; FILE REFERENCE: ABCA5, 6, 9, 10
;; CURRENT APPLICATION NUMBER: US/10/005,338B
;; CURRENT FILING DATE: 2001-12-07
;; PRIOR APPLICATION NUMBER: US 60/263,231
;; PRIOR FILING DATE: 2001-01-23
;; PRIOR APPLICATION NUMBER: FR 00403440.1
;; PRIOR FILING DATE: 2000-12-07
;; NUMBER OF SEQ ID NOS: 217
;; SOFTWARE: Patent in Ver. 2.1
;; SEQ ID NO 36
;; LENGTH: 92
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-005-338B-36

Query Match 1.2%; Score 21.8; DB 14; Length 92;
Best Local Similarity 56.2%; Pred. No. 7.6e+04;
Matches 41; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 1349 TTAATTTCATGACCTTTTGCTGAGTCTGCTTTATAGCAGATTTCTTTTGAAG 1408
DB 91 TGAACACAGGATCTTTTCGCACTTTCTTTCCACATTTTAGTTCATCATCTTAGAA 32
QY 1409 AAAATTCAATAG 1421
DB 31 CAAACACGAATAG 19

THIS PAGE BLANK (uspro)